Diseño y análisis de experimentos

Montgomery

Segunda edición

LIMUSA WILEY

DISEÑO Y ANÁLISIS DE EXPERIMENTOS

DISEÑO Y ANÁLISIS DE EXPERIMENTOS

segunda edición

Douglas C. Montgomery
UNIVERSIDAD ESTATAL DE ARIZONA

LIMUSA WILEY

VERSIÓN AUTORIZADA EN ESPAÑOL DE LA OBRA PUBLICADA EN INGLÉS CON EL TÍTULO:
DESIGN AND ANALYSIS OF EXPERIMENTS
© JOHN WILEY & SONS, INC., NEW YORK, CHICHESTER, BRISBANE, SINGAPORE, TORONTO AND WEINHEIM.

COLABORADOR EN LA TRADUCCIÓN:
RODOLFO PIÑA GARCÍA

REVISIÓN:
GRISELDA ZETINA VÉLEZ
INGENIERA QUÍMICA POR LA FACULTAD DE QUÍMICA DE LA UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO. DOCENTE EN MATEMÁTICAS. PROFESORA EN LA ESCUELA DE CIENCIAS QUÍMICAS DE LA UNIVERSIDAD LA SALLE

GRUPO NORIEGA EDITORES
BALDERAS 95, MÉXICO, D.F.
C.P. 06040
8503 8050
01(800) 706 9100
5512 2903
limusa@noriega.com.mx
www.noriega.com.mx

CANIEM NÚM. 121

HECHO EN MÉXICO
ISBN 968-18-6156-6
3.2



Prefacio

El presente libro es un texto de introducción que aborda el diseño y análisis de experimentos. Tiene como base los cursos sobre diseño de experimentos que he impartido durante más de 25 años en la Universidad Estatal de Arizona, la Universidad de Washington y el Instituto de Tecnología de Georgia. Refleja asimismo los métodos que he encontrado útiles en mi propia práctica profesional como consultor en ingeniería y estadística en las áreas generales de diseño de productos y procesos, mejoramiento de procesos e ingeniería de control de calidad.

El libro está destinado a estudiantes que han llevado un primer curso de métodos estadísticos. Este curso previo debe incluir por lo menos algunas de las técnicas de estadística descriptiva, la distribución normal y una introducción a los conceptos básicos de los intervalos de confianza y la prueba de hipótesis para medias y varianzas. Los capítulos 10 y 11 requieren un manejo elemental de álgebra matricial.

Como los requisitos para llevar este curso son relativamente modestos, este libro puede usarse también en un segundo curso de estadística enfocado en el diseño estadístico de experimentos para estudiantes de licenciatura de ingeniería, física, ciencias físicas y químicas, matemáticas y otros campos de las ciencias. Durante varios años he impartido un curso basado en este libro en el primer año de estudios de posgrado de ingeniería. Los estudiantes de este curso provienen de los campos tradicionales de ingeniería, física, química, matemáticas, investigación de operaciones y estadística. También he usado este libro como base de un curso breve para el sector industrial sobre diseño de experimentos para técnicos en ejercicio con una amplia diversidad en su formación profesional. Se incluyen numerosos ejemplos que ilustran todas las técnicas de diseño y análisis. Estos ejemplos se basan en aplicaciones del diseño experimental en el mundo real, y se han tomado de diferentes campos de la ingeniería y las ciencias. Esto lleva al terreno de las aplicaciones a un curso académico para ingenieros y científicos y hace de este libro una útil herramienta de referencia para experimentadores de una amplia gama de disciplinas.

ACERCA DEL LIBRO

La presente edición constituye una revisión sustancial del libro. He procurado mantener el equilibrio entre los tópicos de diseño y análisis; sin embargo, hay varios temas y ejemplos nuevos; asimismo he reorganizado gran parte del material. En la presente edición se resalta más el uso de la computadora. Durante los últimos años han surgido varios productos de software excelentes que auxilian al experimentador en las fases del diseño y el análisis para esta materia. He incluido las salidas de dos de estos productos, Minitab y *Design-Expert,* en varias partes del texto. Minitab es un paquete de software de estadística de carácter general ampliamente disponible, que cuenta con útiles herramientas de análisis de datos y que maneja bastante bien el análisis de experimentos tanto con factores fijos como aleatorios (incluyendo el modelo mixto). *Design-Expert* es un paquete que se enfoca exclusivamente en el diseño experimental. Tiene muchas herramientas para la construcción y evaluación de diseños, así como múltiples características de análisis. En el sitio *web* de este libro puede obtenerse una versión para estudiantes de *Design-Expert*, y se hace una amplia recomendación para usarlo. Exhorto a todos los profesores que usen este libro para que incorporen software de computadora en sus cursos. En mi caso particular, llevo a todas mis clases una compu-

tadora *laptop* y un monitor, y todos los diseños o tópicos del análisis tratados en clase se ilustran con la computadora.

En esta edición destaco aún más la conexión entre el experimento y el modelo que puede desarrollar el experimentador a partir de los resultados del experimento. Los ingenieros (y en gran medida los científicos de la física y la química) aprenden los mecanismos físicos y sus modelos mecanicistas fundamentales al principio de su formación académica, pero en la mayor parte de sus carreras profesionales tendrán que trabajar con estos modelos. Los experimentos diseñados estadísticamente ofrecen al ingeniero una base válida para desarrollar un *modelo empírico* del sistema bajo estudio. Después este modelo empírico puede manipularse (tal vez utilizando una superficie de respuesta o una gráfica de contorno, o quizá matemáticamente) como cualquier otro modelo de ingeniería. A lo largo de muchos años de docencia he descubierto que este enfoque es muy eficaz para despertar el entusiasmo por los experimentos diseñados estadísticamente en la comunidad de ingeniería. En consecuencia, al inicio del libro planteo la noción de un modelo empírico fundamental para el experimento y las superficies de respuesta y destaco la importancia del mismo.

También me he esforzado por presentar mucho más rápido los puntos críticos en los que intervienen los diseños factoriales. Para facilitar este objetivo, condensé en un solo capítulo (el 3) el material introductorio sobre los experimentos completamente aleatorizados con un solo factor y el análisis de varianza. He ampliado el material sobre los diseños factoriales y factoriales fraccionados (capítulos 5 al 9) en un esfuerzo por hacer que el material fluya con mayor eficiencia en la perspectiva tanto del lector como del profesor y por hacer mayor hincapié en el modelo empírico. El capítulo sobre las superficies de respuesta (el 11) sigue inmediatamente al material sobre diseños factoriales y factoriales fraccionados y modelado de regresiones. He ampliado este capítulo, agregando nuevo material sobre diseños óptimos alfabéticos, experimentos con mezclas y el problema de un diseño paramétrico robusto. En los capítulos 12 y 13 se analizan experimentos que incluyen efectos aleatorios, así como algunas aplicaciones de estos conceptos en diseños anidados y parcelas subdivididas. El capítulo 14 es una descripción general de temas importantes de diseño y análisis: la respuesta no normal, el método de Box-Cox para seleccionar la forma de una transformación, y otras alternativas; experimentos factoriales no balanceados; el análisis de covarianza, incluyendo covariables en un diseño factorial y mediciones repetidas.

A lo largo del libro he destacado la importancia del diseño experimental como una herramienta que el ingeniero en ejercicio puede usar en el diseño y desarrollo de productos, así como en el desarrollo y mejoramiento de procesos. Se ilustra el uso del diseño experimental en el desarrollo de productos que sean robustos a factores ambientales y a otras fuentes de variabilidad. Considero que el uso del diseño experimental en las fases iniciales del ciclo de un producto puede reducir sustancialmente el tiempo y el costo de conducirlo, redundando en procesos y productos con un mejor desempeño en campo y una mayor confiabilidad que los que se desarrollan utilizando otros enfoques.

El libro contiene más material del que puede cubrirse sin prisas en un solo curso, por lo que espero que los profesores puedan variar el contenido de cada curso o bien estudiar más a fondo algunos temas, dependiendo de los intereses de la clase. Al final de cada capítulo hay un grupo de problemas (excepto en el 1). El alcance de estos problemas varía desde ejercicios de cálculo, destinados a consolidar los fundamentos, hasta la ampliación de principios básicos.

Mi curso en la universidad lo enfoco principalmente en los diseños factoriales y factoriales fraccionados. En consecuencia, por lo general cubro el capítulo 1, el capítulo 2 (muy rápido), la mayor parte del capítulo 3, el capítulo 4 (sin incluir el material sobre bloques incompletos y mencionando sólo brevemente los cuadrados latinos), y trato en detalle los capítulos 5 al 8 sobre diseños factoriales con dos niveles y diseños factoriales fraccionados. Para concluir el curso, introduzco la metodología de superficies de respuesta (capítulo 11) y hago un repaso general de los modelos con efectos aleatorios (capítulo 12) y los diseños anidados y en parcelas subdivididas (capítulo 13). Siempre pido a los estudiantes que realicen un

proyecto semestral que consiste en diseñar, conducir y presentar los resultados de un experimento diseñado estadísticamente. Les pido que trabajen en equipos, pues es la manera en que se realiza la mayor parte de la experimentación industrial. Deben hacer la presentación de los resultados de su proyecto de manera oral y por escrito.

MATERIAL SUPLEMENTARIO DEL TEXTO

Con esta edición he preparado un suplemento para cada capítulo del libro. En este material suplementario se desarrollan temas que no pudieron tratarse con mayor detalle en el libro. También presento algunos temas que no aparecen expresamente en el libro, pero que para algunos estudiantes y profesionistas en ejercicio podría resultar de utilidad una introducción de los mismos. El nivel matemático de parte de este material es más elevado que el del texto. Estoy consciente de que los profesores usan este libro con una amplia variedad de audiencias, y es posible que algunos cursos de diseño más avanzados puedan beneficiarse al incluir varios de los temas del material suplementario del texto. Este material está en formato electrónico en el CD/ROM del profesor (disponible sólo en inglés) y se encuentra en el sitio *web* de este libro.

SITIO WEB

En el sitio *web* *http://www.wiley.com/legacy/college/engin/montgomery316490/student/student.html* está disponible el material de apoyo para profesores y estudiantes. Este sitio se usará para comunicar información acerca de innovaciones y recomendaciones para el uso eficaz de este texto. El material suplementario del texto puede encontrarse en este sitio, junto con versiones electrónicas de las series de datos utilizadas en los ejemplos y los problemas de tarea, un plan de estudios del curso y proyectos semestrales del curso en la Universidad Estatal de Arizona.

RECONOCIMIENTOS

Expreso mi agradecimiento a los muchos estudiantes, profesores y colegas que han usado antes este libro y quienes me han hecho llegar útiles sugerencias para esta revisión. Las contribuciones de los doctores Raymond H. Myers, G. Geoffrey Vining, Dennis Lin, John Ramberg, Joseph Pignatiello, Lloyd S. Nelson, Andre Khuri, Peter Nelson, John A. Cornell, George C. Runger, Bert Keats, Dwayne Rollier, Norma Hubele, Cynthia Lowry, Russell G. Heikes, Harrison M. Wadsworth, William W. Hines, Arvind Shah, Jane Ammons, Diane Schaub, Pat Spagon y William DuMouche, y los señores Mark Anderson y Pat Whitcomb fueron particularmente invaluables. Mi Jefe de Departamento, el doctor Gary Hogg, ha proporcionado un ambiente intelectualmente estimulante en el cual trabajar.

Las contribuciones de los profesionistas en activo con quienes he trabajado han sido invaluables. Es imposible mencionarlos a todos, pero algunos de los principales son Dan McCarville y Lisa Custer de Motorola; Richard Post de Intel; Tom Bingham, Dick Vaughn, Julián Anderson, Richard Alkire y Chase Neilson de Boeing Company; Mike Goza, Don Walton, Karen Madison, Jeff Stevens y Bob Kohm de Alcoa; Jay Gardiner, John Butora, Dana Lesher, Lolly Marwah, Paul Tobias y Leon Mason de IBM; Elizabeth A. Peck de The Coca-Cola Company; Sadri Khalessi y Franz Wagner de Signetics; Robert V. Baxley de Monsanto Chemicals; Harry Peterson-Nedry y Russell Boyles de Precision Castparts Corporation; Bill New y Randy Schmid de Allied-Signal Aerospace; John M. Fluke, hijo, de John Fluke Manufacturing

Company; Larry Newton y Kip Howlett de Georgia-Pacific, y Ernesto Ramos de BBN Software Products Corporation.

Me encuentro en deuda con el profesor E.S. Pearson y con *Biometrika*, John Wiley & Sons, Prentice-Hall, The American Statistical Association, The Institute of Mathematical Statistics y los editores de *Biometrics* por el permiso para usar material protegido por derechos de autor. Lisa Custer realizó un excelente trabajo de presentación de las soluciones que aparecen en el CD/ROM del profesor, y la doctora Cheryl Jennings realizó una corrección de estilo eficaz y de suma utilidad. Estoy agradecido con la Office of Naval Research, la National Science Foundation, las compañías integrantes de NSF/Industry/University Cooperative Research Center in Quality and Reliability Engineering de la Universidad Estatal de Arizona, e IBM Corporation por apoyar gran parte de mis investigaciones de estadística y diseño experimental de ingeniería.

Douglas C. Montgomery
Tempe, Arizona

Contenido

1

Introducción

1-1 ESTRATEGIA DE EXPERIMENTACIÓN

Investigadores de prácticamente todos los campos de estudio llevan a cabo experimentos, por lo general para descubrir algo acerca de un proceso o sistema particular. En un sentido literal, un experimento es una **prueba**. En una perspectiva más formal, un **experimento** puede definirse como una prueba o serie de pruebas en las que se hacen cambios deliberados en las variables de entrada de un proceso o sistema para observar e identificar las razones de los cambios que pudieran observarse en la respuesta de salida.

Este libro trata de la planeación y realización de experimentos y del análisis de los datos resultantes a fin de obtener conclusiones válidas y objetivas. La atención se centra en los experimentos de ingeniería y las ciencias físicas y químicas. En ingeniería, la experimentación desempeña un papel importante en el diseño de productos nuevos, el desarrollo de procesos de manufactura y el mejoramiento de procesos. El objetivo en muchos casos sería desarrollar un proceso **robusto**, es decir, un proceso que sea afectado en forma mínima por fuentes de variabilidad externas.

Como ejemplo de un experimento, suponga que un ingeniero metalúrgico tiene interés en estudiar el efecto de dos procesos diferentes de endurecimiento, el templado en aceite y el templado en agua salada, sobre una aleación de aluminio. El objetivo del experimentador es determinar cuál de las dos soluciones de templado produce la dureza máxima para esta aleación particular. El ingeniero decide someter varios ejemplares o muestras para ensayo de la aleación a cada medio de templado y medir la dureza de los ejemplares después del templado. Para determinar cuál de las soluciones es la mejor, se usará la dureza promedio de los ejemplares tratados en cada solución de templado.

Al examinar este sencillo experimento salen a relucir varias cuestiones importantes:

1. ¿Estas dos soluciones son los únicos medios de templado de interés potencial?
2. ¿Hay en este experimento otros factores que podrían afectar la dureza y que deberían investigarse o controlarse?
3. ¿Cuántas muestras para ensayo de la aleación deberán probarse en cada solución de templado?
4. ¿Cómo deberán asignarse las muestras para ensayo de prueba a las soluciones de templado y en qué orden deberán colectarse los datos?

5. ¿Qué método de análisis de datos deberá usarse?
6. ¿Qué diferencia en la dureza promedio observada entre los dos medios de templado se considerará importante?

Todas estas preguntas, y tal vez muchas más, tendrán que responderse satisfactoriamente antes de llevar a cabo el experimento.

En cualquier experimento, los resultados y las conclusiones que puedan sacarse dependen en gran medida de la manera en que se recabaron los datos. Para ilustrar este punto, suponga que el ingeniero metalúrgico del experimento anterior utilizó ejemplares de una hornada para el templado en aceite y ejemplares de una segunda hornada para el templado en agua salada. Entonces, cuando compare la dureza promedio, el ingeniero no podrá saber qué parte de la diferencia observada es resultado de la solución de templado y qué parte es el resultado de diferencias inherentes entre las hornadas.[1] Por lo tanto, el método utilizado para recabar los datos ha afectado de manera adversa las conclusiones que pueden sacarse del experimento.

En general, los experimentos se usan para estudiar el desempeño de procesos y sistemas. El proceso o sistema puede representarse con el modelo ilustrado en la figura 1-1. El proceso puede por lo general visualizarse como una combinación de máquinas, métodos, personas u otros recursos que transforman cierta entrada (con frecuencia un material) en una salida que tiene una o más respuestas observables. Algunas variables del proceso $x_1, x_2, ..., x_p$ son controlables, mientras que otras $z_1, z_2, ..., z_q$ son no controlables (aunque pueden serlo para los fines de una prueba). Los objetivos del experimento podrían comprender los siguientes:

1. Determinar cuáles son las variables que tienen mayor influencia sobre la respuesta y.
2. Determinar cuál es el ajuste de las x que tiene mayor influencia para que y esté casi siempre cerca del valor nominal deseado.
3. Determinar cuál es el ajuste de las x que tiene mayor influencia para que la variabilidad de y sea reducida.
4. Determinar cuál es el ajuste de las x que tiene mayor influencia para que los efectos de las variables no controlables $z_1, z_2, ..., z_q$ sean mínimos.

Como se puede ver por el análisis anterior, los experimentos incluyen muchas veces varios factores. Habitualmente, uno de los objetivos de la persona que realiza un experimento, llamada el **experimentador**, es determinar la influencia que tienen estos factores sobre la respuesta de salida del sistema. Al enfo-

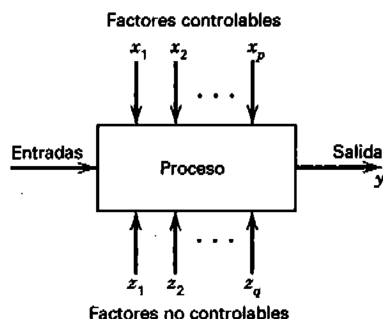


Figura 1-1 Modelo general de un proceso o sistema.

[1] Un especialista en diseño experimental diría que los efectos de los medios de templado y las hornadas *se confundieron*; es decir, los efectos de estos dos factores no pueden separarse.

que general para planear y llevar a cabo el experimento se le llama **estrategia de experimentación**. Existen varias estrategias que podría usar un experimentador. Se ilustrarán algunas de ellas con un ejemplo muy sencillo.

Al autor le gusta mucho jugar golf. Desafortunadamente, no le agrada practicar, por lo que siempre busca la manera más sencilla para bajar su puntuación. Algunos de los factores que él considera importantes, o que podrían influir en su puntuación, son los siguientes:

1. El tipo de palo usado (grande o normal).
2. El tipo de pelota usada (de goma de balata o de tres piezas).
3. Caminar cargando los palos de golf o hacer el recorrido en un carrito.
4. Beber agua o cerveza durante el juego.
5. Jugar en la mañana o en la tarde.
6. Jugar cuando hace frío o cuando hace calor.
7. El tipo de spikes usados en los zapatos de golf (metálicos o de hule).
8. Jugar en un día con viento o en uno apacible.

Evidentemente, hay muchos otros factores que podrían considerarse, pero supongamos que éstos son los de interés primario. Además, teniendo en cuenta su larga experiencia en el juego, el autor decide que los factores 5 al 8 pueden ignorarse; es decir, estos factores no son importantes porque sus efectos son tan pequeños que carecen de valor práctico. Los ingenieros y los investigadores deben tomar a menudo este tipo de decisiones acerca de algunos de los factores que examinan en experimentos reales.

Consideremos ahora cómo podrían probarse experimentalmente los factores 1 al 4 para determinar su efecto sobre la puntuación del autor. Suponga que en el curso del experimento pueden jugarse un máximo de ocho rondas de golf. Un enfoque consistiría en seleccionar una combinación arbitraria de estos factores, probarlos y ver qué ocurre. Por ejemplo, suponga que se selecciona la combinación del palo grande, la pelota de goma de balata, el carrito y el agua, y que la puntuación resultante es 87. Sin embargo, durante la ronda el autor notó varios tiros descontrolados con el palo grande (en el golf, grande no siempre es sinónimo de bueno) y, en consecuencia, decide jugar otra ronda con el palo normal, manteniendo los demás factores en los mismos niveles usados anteriormente. Este enfoque podría continuar de manera casi indefinida, cambiando los niveles de uno (o quizá dos) de los factores para la prueba siguiente, con base en el resultado de la prueba en curso. Esta estrategia de experimentación, conocida como **enfoque de la mejor conjetura**, es común entre ingenieros y científicos. Funciona de manera adecuada si los experimentadores cuentan con una gran cantidad de conocimientos técnicos o teóricos del sistema que están estudiando, así como amplia experiencia práctica. Sin embargo, el enfoque de la mejor conjetura presenta al menos dos desventajas. Primera, supóngase que la mejor conjetura inicial no produce los resultados deseados. Entonces el experimentador tiene que hacer otra conjetura acerca de la combinación correcta de los niveles de los factores. Esto podría continuar por mucho tiempo, sin garantía alguna de éxito. Segunda, supóngase que la mejor conjetura inicial produce un resultado satisfactorio. Entonces, el experimentador se ve tentado a suspender las pruebas, aun cuando no hay ninguna garantía de que se ha encontrado la mejor solución.

Otra estrategia de experimentación muy común en la práctica es el **enfoque de un factor a la vez**. Este método consiste en seleccionar un punto de partida, o línea **base** de los niveles, para cada factor, para después variar sucesivamente cada factor en su rango, manteniendo constantes los factores restantes en el nivel base. Después de haber realizado todas las pruebas, se construye por lo general una serie de gráficas en las que se muestra la forma en que la variable de respuesta es afectada al variar cada factor manteniendo los demás factores constantes. En la figura 1-2 se presenta una serie de gráficas para el experimento del golf, utilizando como línea base los niveles de los cuatro factores: el palo grande, la pelota de goma de

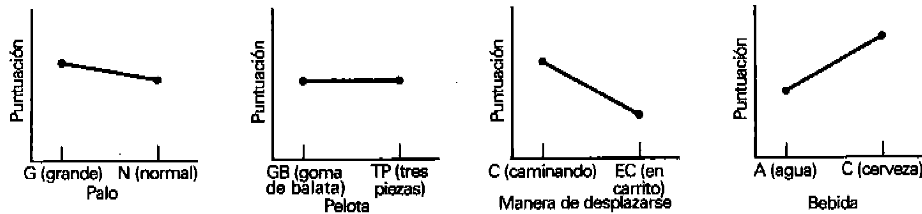


Figura 1-2 Resultados de la estrategia de un factor a la vez para el experimento de golf.

balata, caminar y beber agua. La interpretación de esta gráfica es directa; por ejemplo, debido a que la pendiente de la curva de la manera de desplazarse es negativa, se concluiría que hacer el recorrido en el carrito mejora la puntuación. Con base en estas gráficas de un factor a la vez, la combinación óptima que se seleccionaría sería el palo normal, desplazarse en el carrito y beber agua. El tipo de pelota de golf aparentemente carece de importancia.

La desventaja principal de la estrategia de un factor a la vez es que no puede tomar en consideración cualquier posible **interacción** entre los factores. Hay una interacción cuando uno de los factores no produce el mismo efecto en la respuesta con niveles diferentes de otro factor. En la figura 1-3 se muestra una interacción entre los factores del tipo de palo y la bebida para el experimento del golf. Observe que si el autor utiliza el palo normal, el tipo de bebida consumida prácticamente no tiene efecto alguno sobre su puntuación, pero si utiliza el palo grande, se obtienen resultados mucho mejores cuando bebe agua en lugar de cerveza. Las interacciones entre factores son muy comunes y, en caso de existir, la estrategia de un factor a la vez casi siempre producirá resultados deficientes. Muchas personas no perciben esto y, en consecuencia, los experimentos de un factor a la vez son comunes en la práctica. (De hecho, algunas personas piensan que esta estrategia se relaciona con el método científico o que es un principio "sólido" de ingeniería.) Los experimentos de un factor a la vez siempre son menos eficientes que otros métodos basados en un enfoque estadístico del diseño experimental. El tema se analizará con mayor detalle en el capítulo 5.

El enfoque correcto para trabajar con varios factores es conducir un experimento **factorial**. Se trata de una estrategia experimental en la que los factores se hacen variar en *conjunto*, en lugar de uno a la vez.

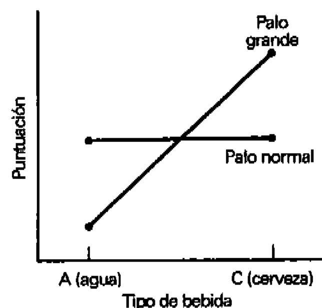


Figura 1-3 Interacción entre el tipo de palo y el tipo de bebida para el experimento del golf.

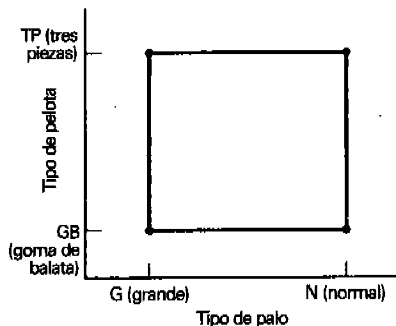


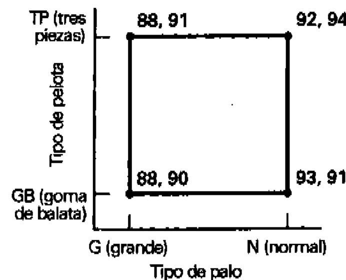
Figura 1-4 Experimento factorial de dos factores que incluye el tipo de palo y el tipo de pelota.

El concepto de diseño experimental factorial es de suma importancia, y varios capítulos de este libro se dedican a presentar experimentos factoriales básicos, así como algunas variantes y casos especiales útiles.

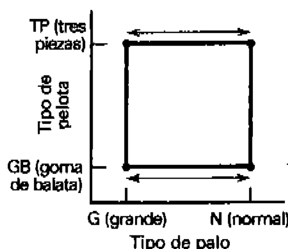
Para ilustrar la forma en que se lleva a cabo un experimento factorial, considere el experimento de golf y suponga que sólo dos de los factores son de interés, el tipo de palo y el tipo de pelota. En la figura 1-4 se muestra un experimento factorial para estudiar los efectos conjuntos de estos dos factores sobre la puntuación de golf del autor. Observe que en este experimento factorial ambos factores tienen dos niveles y que en el diseño se usan todas las combinaciones posibles de los niveles de ambos factores. Geométricamente, las cuatro corridas forman los vértices de un cuadrado. A este tipo particular de experimento factorial se le llama **diseño factorial 2^2** (dos factores, cada uno con dos niveles). Debido a que el autor considera razonable suponer que jugará ocho rondas de golf para investigar estos factores, un plan factible sería jugar dos rondas de golf con cada combinación de los niveles de los factores, como se muestra en la figura 1-4. Un diseñador de experimentos diría que se han hecho dos **réplicas** del diseño. Este diseño experimental permitiría al experimentador investigar los efectos individuales (o los efectos **principales**) de cada factor y determinar si existe alguna interacción entre los factores.

En la figura 1-5*a* se presentan los resultados obtenidos al realizar el experimento factorial de la figura 1-4. En los vértices del cuadrado se indican las puntuaciones de cada ronda de golf jugada con las cuatro combinaciones de prueba. Observe que hay cuatro rondas de golf que proporcionan información acerca del uso del palo normal y cuatro rondas que proporcionan información sobre el uso del palo grande. Al encontrar la diferencia promedio de las puntuaciones que están en los lados derecho e izquierdo del cuadrado (como en la figura 1-5*b*), se tiene una medida del efecto de cambiar del palo grande al palo normal, o

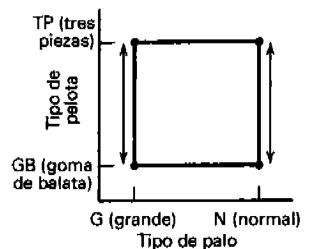
$$\text{Efecto del palo} = \frac{92+94+93+91}{4} - \frac{88+91+88+90}{4} = 3.25$$



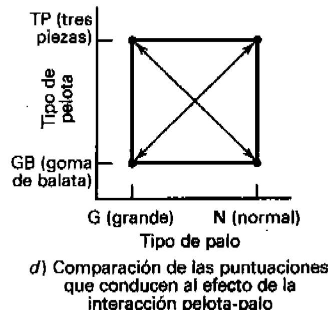
a) Puntuaciones del experimento de golf



b) Comparación de las puntuaciones que conducen al efecto del palo



c) Comparación de las puntuaciones que conducen al efecto de la pelota



d) Comparación de las puntuaciones que conducen al efecto de la interacción pelota-palo

Figura 1-5 Puntuaciones del experimento del golf de la figura 1-4 y cálculo de los efectos de los factores.

Es decir, en promedio, al cambiar del palo grande al normal la puntuación se incrementa 3.25 golpes por ronda. De manera similar, la diferencia promedio de las cuatro puntuaciones de la parte superior del cuadrado y de las cuatro puntuaciones de la parte inferior miden el efecto del tipo de pelota usado (ver la figura 1-5*c*):

$$\text{Efecto de la pelota} = \frac{88+91+92+94}{4} - \frac{88+90+93+91}{4} = 0.75$$

Por último, puede obtenerse una medida del efecto de la interacción entre el tipo de pelota y el tipo de palo restando la puntuación promedio en la diagonal de izquierda a derecha del cuadrado de la puntuación promedio de la diagonal de derecha a izquierda (ver la figura 1-5*d*), cuyo resultado es

$$\text{Efecto de la interacción pelota-palo} = \frac{92+94+88+90}{4} - \frac{88+91+93+91}{4} = 0.25$$

Los resultados de este experimento factorial indican que el efecto del palo es mayor que el efecto de la pelota o que el de la interacción. Podrían usarse pruebas estadísticas para determinar si cualquiera de estos efectos difiere de cero. De hecho, el caso es que hay evidencia estadística razonablemente sólida de que el efecto del palo difiere de cero y de que no es el caso para los otros dos efectos. Por lo tanto, tal vez el autor debería jugar siempre con el palo grande.

En este sencillo ejemplo se pone de manifiesto una característica muy importante del experimento factorial: en los diseños factoriales se hace el uso más eficiente de los datos experimentales. Note que este experimento incluyó ocho observaciones, y que las ocho observaciones se usan para calcular los efectos del palo, de la pelota y de la interacción. Ninguna otra estrategia de experimentación hace un uso tan eficiente de los datos. Ésta es una característica importante y útil de los diseños factoriales.

El concepto de experimento factorial puede extenderse a tres factores. Suponga que el autor desea estudiar los efectos del tipo de palo, el tipo de pelota y el tipo de bebida consumida sobre su puntuación de golf. Suponiendo que los tres factores tienen dos niveles, puede establecerse un diseño factorial como el que se muestra en la figura 1-6. Observe que hay ocho combinaciones de prueba de estos tres factores con los dos niveles de cada uno de ellos y que estos ocho ensayos pueden representarse geométricamente como los vértices de un cubo. Se trata de un ejemplo de un **diseño factorial 2^3**. Como el autor sólo desea jugar ocho rondas de golf, este experimento requeriría que se juegue una ronda con cada una de las combinaciones de los factores representadas por los ocho vértices del cubo de la figura 1-6. Sin embargo, al comparar esta situación con el diseño factorial de dos factores de la figura 1-4, el diseño factorial 2^3 produciría la misma información acerca de los efectos de los factores. Por ejemplo, en ambos diseños hay cuatro pruebas que proporcionan información acerca del palo normal y cuatro pruebas que proporcionan

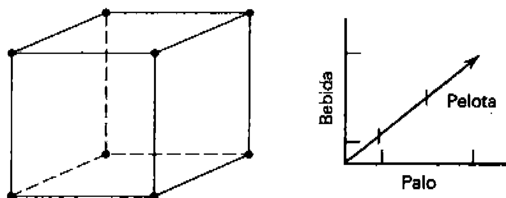


Figura 1-6 Experimento factorial de tres factores que incluye el tipo de palo, el tipo de pelota y el tipo de bebida.

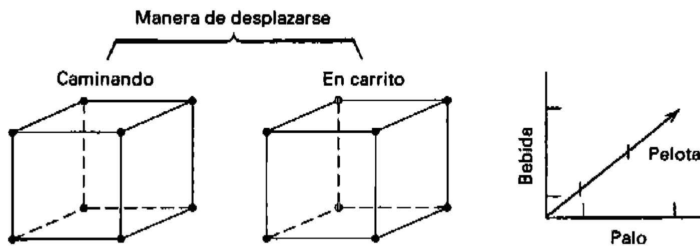


Figura 1-7 Experimento factorial de cuatro factores que incluye el tipo de palo, el tipo de pelota, el tipo de bebida y la manera de desplazarse.

información acerca del palo grande, suponiendo que se repite dos veces cada corrida del diseño de dos factores de la figura 1-4.

En la figura 1-7 se ilustra la forma en que podrían investigarse los cuatro factores —el palo, la pelota, la bebida y la manera de desplazarse (caminando o en carrito)— en un **diseño factorial 2^4**. Como en cualquier diseño factorial, se usan todas las combinaciones posibles de los niveles de los factores. Puesto que los cuatro factores tienen dos niveles, sigue siendo posible hacer la representación geométrica de este diseño experimental mediante un cubo (en realidad un hipercubo).

En general, si hay k factores, cada uno con dos niveles, el diseño factorial requeriría 2^k corridas. Por ejemplo, el experimento de la figura 1-7 requiere 16 corridas. Evidentemente, cuando el número de factores de interés aumenta, el número de corridas requeridas se incrementa con rapidez; por ejemplo, un experimento con 10 factores en el que todos los factores tienen dos niveles requeriría 1024 corridas. Esto pronto se vuelve impracticable en lo que se refiere al tiempo y los recursos. En el experimento del golf, el autor sólo puede jugar ocho rondas, por lo que incluso el experimento de la figura 1-7 resulta demasiado largo.

Por fortuna, cuando se trabaja con cuatro, cinco o más factores, por lo general no es necesario probar todas las combinaciones posibles de los niveles de los factores. Un **experimento factorial fraccionado** es una variación del diseño factorial básico en la que sólo se realiza un subconjunto de las corridas. En la figura 1-8 se ilustra un diseño factorial fraccionado para la versión de cuatro factores del experimento del golf. Este diseño requiere sólo 8 corridas en lugar de las 16 originales y se llamaría **fracción un medio**. Si el autor sólo puede jugar ocho rondas de golf, éste es un excelente diseño en el cual estudiar los cuatro factores. Proporcionará información adecuada acerca de los efectos principales de los cuatro factores, así como cierta información acerca de la forma en que interactúan estos factores.

Los diseños factoriales fraccionados son muy comunes en la investigación y el desarrollo industrial, así como en el mejoramiento de procesos. Estos diseños se analizarán en el capítulo 8.

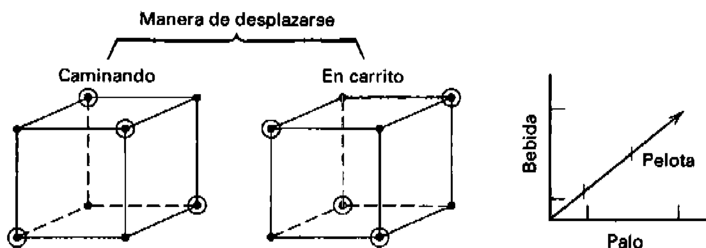


Figura 1-8 Experimento factorial fraccionado de cuatro factores que incluye el tipo de palo, el tipo de pelota, el tipo de bebida y la manera de desplazarse.

1-2 ALGUNAS APLICACIONES TÍPICAS DEL DISEÑO EXPERIMENTAL

Los métodos del diseño experimental han encontrado amplia aplicación en diversas disciplinas. De hecho, la experimentación puede considerarse parte del proceso científico y uno de los medios para conocer el funcionamiento de sistemas y procesos. En general, el aprendizaje ocurre a través de una serie de actividades en las que se hacen conjeturas acerca de un proceso, se llevan a cabo experimentos para generar datos del proceso y después se usa la información del experimento para establecer nuevas conjeturas, lo que lleva a nuevos experimentos, y así sucesivamente.

El diseño experimental es una herramienta de importancia fundamental en el ámbito de la ingeniería para mejorar el desempeño de un proceso de manufactura. También tiene múltiples aplicaciones en el desarrollo de procesos nuevos. La aplicación de las técnicas del diseño experimental en las fases iniciales del desarrollo de un proceso puede redundar en

1. Mejoras en el rendimiento del proceso.
2. Variabilidad reducida y conformidad más cercana con los requerimientos nominales o proyectados.
3. Reducción del tiempo de desarrollo.
4. Reducción de los costos globales.

Los métodos del diseño experimental desempeñan también un papel importante en las actividades del **diseño de ingeniería,** donde se desarrollan productos nuevos y se hacen mejoramientos en los productos existentes. Entre las aplicaciones del diseño experimental en el diseño de ingeniería se encuentran:

1. La evaluación y comparación de configuraciones de diseños básicos.
2. La evaluación de materiales alternativos.
3. La selección de los parámetros del diseño para que el producto tenga un buen funcionamiento en una amplia variedad de condiciones de campo, es decir, para que el producto sea **robusto**.
4. La determinación de los parámetros clave del diseño del producto que afectan el desempeño del mismo.

El uso del diseño experimental en estas áreas puede redundar en productos cuya fabricación sea más sencilla, en productos que tengan un desempeño y confiabilidad de campo mejorados, en costos de producción más bajos y en tiempos más cortos para el diseño y desarrollo del producto. A continuación se presentan varios ejemplos que ilustran algunas de estas ideas.

EJEMPLO 1-1 ···

Caracterización de un proceso

En el proceso de fabricación de tarjetas de circuitos impresos se utiliza una máquina de soldadura líquida. La máquina limpia las tarjetas en un fundente, las somete a un proceso de precalentamiento y después las hace pasar por una onda de soldadura líquida mediante una transportadora. En este proceso de soldadura se hacen las conexiones eléctricas y mecánicas de los componentes recubiertos de plomo en la tarjeta.

El proceso opera actualmente con un nivel de defectos aproximado de 1%. Es decir, cerca de 1% de las juntas de soldadura de una tarjeta son defectuosas y requieren corrección manual. Sin embargo, debido a que la tarjeta de circuitos impresos promedio contiene más de 2000 juntas de soldadura, incluso un nivel de defectos de 1% representa un número demasiado alto de juntas de soldadura que requieren corrección. Al ingeniero responsable del proceso en esta área le gustaría usar un experimento diseñado para determinar cuáles son los parámetros de la máquina que influyen en la ocurrencia de los defectos de soldadura y qué ajustes deberían hacerse en dichas variables para reducir los defectos de soldadura.

En la máquina de soldadura líquida hay diversas variables que pueden controlarse. Éstas incluyen:

1. La temperatura de la soldadura.
2. La temperatura del precalentamiento.
3. La velocidad de la transportadora.
4. El tipo de fundente.
5. La gravedad específica del fundente.
6. La profundidad de la onda de soldadura.
7. El ángulo de la transportadora.

Además de estos factores controlables, hay otros que no es sencillo manejar durante el proceso de fabricación, aunque podrían controlarse para los fines de una prueba. Éstos son:

1. El espesor de la tarjeta de circuitos impresos.
2. El tipo de componentes usados en la tarjeta.
3. La disposición de los componentes en la tarjeta.
4. El operador.
5. La rapidez de producción.

En esta situación, el interés del ingeniero es **caracterizar** la máquina de soldadura líquida; es decir, quiere determinar los factores (tanto los controlables como los no controlables) que afectan la ocurrencia de defectos en las tarjetas de circuitos impresos. Para ello puede diseñar un experimento que le permitirá estimar la magnitud y dirección de los efectos de los factores; es decir, cuánto cambia la variable de respuesta (defectos por unidad) cuando se modifica cada factor, y si la modificación de los factores en conjunto produce resultados diferentes que los obtenidos mediante el ajuste individual de los factores; es decir, ¿existe interacción entre los factores? En ocasiones a un experimento como éste se le llama **experimento tamiz o de exploración** exhaustiva. De manera típica, los experimentos tamiz incluyen el uso de diseños factoriales fraccionados, como en el ejemplo del golf de la figura 1-8.

La información obtenida de este experimento tamiz se usará para identificar los factores críticos del proceso y determinar la dirección del ajuste de dichos factores a fin de conseguir una reducción adicional del número de defectos por unidad. El experimento también puede proporcionar información acerca de los factores que deberían controlarse con mayor atención durante el proceso de fabricación a fin de evitar los niveles elevados de defectos y el desempeño errático del proceso. Por lo tanto, una consecuencia del experimento podría ser la aplicación de técnicas como las cartas de control a una o más de las **variables del proceso** (la temperatura de la soldadura, por ejemplo), aunadas a las cartas de control de la producción del proceso. Con el tiempo, si se consigue una mejoría sensible del proceso, quizá sea posible basar la mayor parte del programa de control del mismo en el control de las variables de entrada del proceso en lugar de aplicar cartas de control a la producción.

EJEMPLO 1-2

Optimización de un proceso

En un experimento de caracterización, el interés suele centrarse en determinar las variables del proceso que afectan la respuesta. El siguiente paso lógico es la optimización, es decir, determinar la región de los factores importantes que conduzca a la mejor respuesta posible. Por ejemplo, si la respuesta es el rendi-

miento, se buscaría la región del rendimiento máximo, mientras que si la respuesta es la variabilidad de una dimensión crítica del producto, se buscaría una región de variabilidad mínima.

Supongamos que el interés se centra en mejorar el rendimiento de un proceso químico. Por los resultados de un experimento de caracterización se sabe que las dos variables más importantes del proceso que influyen en el rendimiento son la temperatura de operación y el tiempo de reacción. El proceso opera actualmente a 145°F y con 2.1 horas de tiempo de reacción, produciendo rendimientos de cerca de 80%. En la figura 1-9 se muestra una vista desde arriba de la región tiempo-temperatura. En esta gráfica las líneas de rendimiento constante se unen para formar los **contornos** de respuesta, y se muestran las líneas de contorno para rendimientos de 60, 70, 80, 90 y 95 por ciento. Estos contornos son las proyecciones en la región tiempo-temperatura de las secciones transversales de la superficie del rendimiento correspondiente a los rendimientos porcentuales arriba mencionados. A esta superficie se le llama en ocasiones **superficie de respuesta**. El personal del proceso no conoce la verdadera superficie de respuesta de la figura 1-9, por lo que se necesitarán métodos experimentales para optimizar el rendimiento con respecto al tiempo y la temperatura.

Para localizar el rendimiento óptimo, es necesario llevar a cabo un experimento en el que se hagan variar conjuntamente el tiempo y la temperatura, es decir, un experimento factorial. En la figura 1-9 se muestran los resultados de un experimento factorial inicial realizado con dos niveles tanto del tiempo como de la temperatura. Las respuestas que se observan en los cuatro vértices del cuadrado indican que, para incrementar el rendimiento, los cambios deberían hacerse en la dirección general del aumento de la

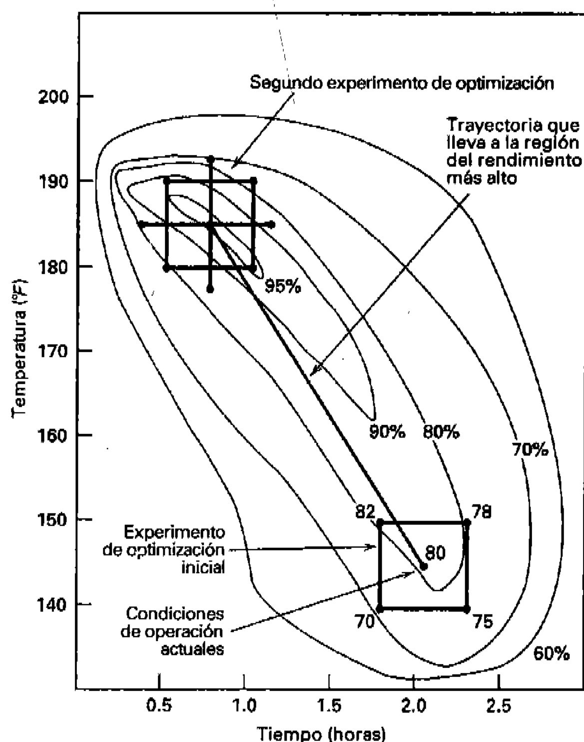


Figura 1-9 Gráfica de contorno del rendimiento como una función del tiempo de reacción y la temperatura de reacción, la cual ilustra la experimentación para optimizar un proceso.

temperatura y la reducción del tiempo de reacción. Se realizarían algunas corridas adicionales en esta dirección, y esta experimentación adicional llevaría a la región del rendimiento máximo.

Una vez que se ha encontrado la región del rendimiento óptimo, el siguiente paso típico sería realizar un segundo experimento. El objetivo de este segundo experimento es desarrollar un modelo empírico del proceso y obtener una estimación más precisa de las condiciones de operación óptimas para el tiempo y la temperatura. A este enfoque para la optimización de un proceso se le llama la **metodología de superficies de respuesta**, la cual se examina en detalle en el capítulo 11. El segundo diseño ilustrado en la figura 1-9 es un **diseño central compuesto**, uno de los diseños experimentales más importantes que se usan en los estudios de optimización de procesos.

EJEMPLO 1-3

Ilustración del diseño de un producto

Con frecuencia los métodos de diseño experimental pueden aplicarse en el proceso de diseño de un producto. Para ilustrar esto, suponga que un grupo de ingenieros está diseñando el gozne de la puerta de un automóvil. La característica de calidad del producto que les interesa es el esfuerzo amortiguador, es decir, la capacidad de retención del tope que impide que la puerta se cierre cuando el vehículo se estaciona en una pendiente. El mecanismo amortiguador consta de un resorte de hojas y un cilindro. Cuando la puerta se abre, el cilindro se desplaza por un arco que hace que el resorte de hojas se comprima. Para cerrar la puerta es necesario vencer la fuerza del resorte, la cual produce el esfuerzo amortiguador. El equipo de ingenieros considera que el esfuerzo amortiguador es una función de los siguientes factores:

1. La distancia que se desplaza el cilindro.
2. La altura del resorte del pivote a la base.
3. La distancia horizontal del pivote al resorte.
4. La altura libre del resorte auxiliar.
5. La altura libre del resorte principal.

Los ingenieros pueden construir un prototipo del mecanismo del gozne en el que es posible variar todos estos factores dentro de ciertos rangos. Una vez que se han identificado los niveles apropiados de estos cinco factores, puede diseñarse un experimento que conste de varias combinaciones de los niveles de los factores, y el prototipo del gozne puede probarse con estas combinaciones. Se obtendrá así información respecto de los factores que tienen una mayor influencia sobre el esfuerzo amortiguador del tope y, mediante el análisis de esta información, podrá mejorarse el diseño del tope.

1-3 PRINCIPIOS BÁSICOS

Si quiere llevarse a cabo un experimento como los descritos en los ejemplos 1-1 al 1-3 con la mayor eficiencia posible, es necesario utilizar un enfoque científico para planearlo. El **diseño estadístico de experimentos** se refiere al proceso para planear el experimento de tal forma que se recaben datos adecuados que puedan analizarse con métodos estadísticos que llevarán a conclusiones válidas y objetivas. El enfoque estadístico del diseño experimental es necesario si se quieren sacar conclusiones significativas de los datos. Cuando el problema incluye datos que están sujetos a errores experimentales, la metodología estadística es el único enfoque objetivo de análisis. Por lo tanto, cualquier problema experimental incluye dos

aspectos: el diseño del experimento y el análisis estadístico de los datos. Estos dos aspectos se encuentran íntimamente relacionados porque el método de análisis depende directamente del diseño empleado. Ambos temas se tratan en este libro.

Los tres principios básicos del diseño experimental son la **realización de réplicas**, la **aleatorización** y la **formación de bloques**. Por realización de réplicas se entiende la repetición del experimento básico. En el experimento metalúrgico analizado en la sección 1-1, una réplica consistiría en el tratamiento de una muestra con el templado en aceite y el tratamiento de una muestra con el templado en agua salada. Por lo tanto, si se tratan cinco ejemplares en cada medio de templado, se dice que se han obtenido cinco **réplicas**. La realización de réplicas posee dos propiedades importantes. Primera, permite al experimentador obtener una estimación del error experimental. Esta estimación del error se convierte en una unidad de medición básica para determinar si las diferencias observadas en los datos son en realidad *estadísticamente* diferentes. Segunda, si se usa la media muestral (por ejemplo, $\bar{y}$) para estimar el efecto de un factor en el experimento, la realización de réplicas permite al experimentador obtener una estimación más precisa de este efecto. Por ejemplo, si σ^2 es la varianza de una observación individual y hay n réplicas, la varianza de la media muestral es

$$\sigma_{\bar{y}}^2 = \frac{\sigma^2}{n}$$

La consecuencia práctica de lo anterior es que si se hicieron $n = 1$ réplicas y se observó $y_1 = 145$ (templado en aceite) y $y_2 = 147$ (templado en agua salada), probablemente no podrán hacerse inferencias satisfactorias acerca del efecto del medio de templado; es decir, la diferencia observada podría ser resultado del error experimental. Por otra parte, si n fue razonablemente grande y el error experimental fue lo suficientemente pequeño, y se observó $\bar{y}_1 < \bar{y}_2$, podría concluirse con una certeza razonable que el templado en agua salada produce una dureza mayor en esta aleación de aluminio particular que el templado en aceite.

Hay una diferencia importante entre réplicas y **mediciones repetidas**. Por ejemplo, suponga que una oblea de silicio se graba con un proceso de grabado químico con plasma para oblea única, y que se hacen tres mediciones de una dimensión crítica de esta oblea. Estas mediciones no son réplicas; son una forma de mediciones repetidas y, en este caso, la variabilidad observada en las tres mediciones repetidas es reflejo directo de la variabilidad inherente del sistema o instrumento de medición. Como otro ejemplo, suponga que, como parte de un experimento en la manufactura de semiconductores, se procesan simultáneamente cuatro obleas en un horno de oxidación con una velocidad del flujo de gas y un tiempo particulares y que se hace después una medición del espesor del óxido en cada oblea. De nueva cuenta, la medición de las cuatro obleas no son réplicas sino mediciones repetidas. En este caso reflejan las diferencias entre las obleas y otras fuentes de variabilidad dentro de esa operación de horneado particular. En las réplicas se reflejan las fuentes de variabilidad tanto **entre** las corridas como (potencialmente) **dentro** de las mismas.

La **aleatorización** es la piedra angular en la que se fundamenta el uso de los métodos estadísticos en el diseño experimental. Por aleatorización se entiende que tanto la asignación del material experimental como el orden en que se realizarán las corridas o ensayos individuales del experimento se determinan al azar. Uno de los requisitos de los métodos estadísticos es que las observaciones (o los errores) sean variables aleatorias con distribuciones independientes. La aleatorización hace por lo general que este supuesto sea válido. La aleatorización correcta del experimento ayuda también a "sacar del promedio" los efectos de factores extraños que pudieran estar presentes. Por ejemplo, suponga que los ejemplares del experimento descrito antes presentan sólo ligeras diferencias en el espesor y que la efectividad del medio de templado puede ser afectado por el espesor del ejemplar. Si todos los ejemplares sometidos al templado en aceite son más gruesos que los sometidos al templado en agua salada, quizá se esté introduciendo un sesgo sistemático en los resultados experimentales. Este sesgo estorba en uno de los medios de templa-

do y, en consecuencia, invalida los resultados obtenidos. Al hacer la asignación aleatoria de los ejemplares al medio de templado este problema se aligera en parte.

Es muy común el uso de programas de computadora para auxiliar a los experimentadores a seleccionar y construir diseños experimentales. Estos programas presentan a menudo las corridas del diseño experimental de manera aleatoria. Por lo general este modo aleatorio se crea utilizando un generador de números aleatorios. Incluso con estos programas de computadora, con frecuencia seguirá siendo necesario que el experimentador haga la asignación del material experimental (como las obleas en los ejemplos de semiconductores mencionados antes), de los operadores, de los instrumentos o herramientas de medición, etc., que se utilizarán en el experimento. Puede recurrirse a tablas de números aleatorios para asegurar que las asignaciones se hacen al azar.

En ocasiones los experimentadores se encuentran con situaciones en las que la aleatorización de un aspecto del experimento es complicada. Por ejemplo, en un proceso químico, la temperatura puede ser una variable muy difícil de modificar, haciendo casi imposible la **aleatorización completa** de este factor. Existen métodos de diseño estadístico para resolver las restricciones sobre la aleatorización. Algunos de estos enfoques se revisarán en capítulos subsecuentes (ver en particular el capítulo 13).

La **formación de bloques** es una técnica de diseño que se utiliza para mejorar la precisión de las comparaciones que se hacen entre los factores de interés. Muchas veces la formación de bloques se emplea para reducir o eliminar la variabilidad transmitida por **factores perturbadores**; es decir, aquellos factores que pueden influir en la respuesta experimental pero en los que no hay un interés específico. Por ejemplo, en un experimento de un proceso químico pueden requerirse dos lotes de materia prima para realizar todas las corridas necesarias. Sin embargo, podría haber diferencias entre los lotes debido a la variabilidad de un proveedor a otro y, en caso de no haber un interés específico en este efecto, los lotes de materia prima se considerarían un factor perturbador. En general, un bloque es un conjunto de condiciones experimentales relativamente homogéneas. En el ejemplo del proceso químico, cada lote de materia prima formaría un bloque, ya que es de esperarse que la variabilidad dentro de un lote sea menor que la variabilidad entre lotes. De manera típica, como en este ejemplo, cada nivel del factor perturbador pasa a ser un bloque. Entonces el experimentador divide las observaciones del diseño estadístico en grupos que se corren en cada bloque. En varias partes del texto se estudia en detalle la formación de bloques, incluyendo los capítulos 4, 5, 7, 8, 9, 11 y 13. En el capítulo 2, sección 2-5.1, se presenta un ejemplo sencillo para ilustrar la estructura básica de la formación de bloques.

Los tres principios básicos del diseño experimental, la aleatorización, la realización de réplicas y la formación de bloques son parte de cada uno de los experimentos. Se ilustrarán y resaltarán repetidamente a lo largo de este libro.

1-4 PAUTAS GENERALES PARA DISEÑAR EXPERIMENTOS

Para aplicar el enfoque estadístico en el diseño y análisis de un experimento, es necesario que todos los que participan en el mismo tengan desde el principio una idea clara de qué es exactamente lo que va a estudiarse, cómo van a colectarse los datos, y al menos una comprensión cualitativa de la forma en que van a analizarse estos datos. En la tabla 1-1 se muestra un esquema general del procedimiento recomendado. A continuación se presenta una breve explicación de este esquema y se elaboran algunos de los puntos clave. Para mayores detalles, ver Coleman y Montgomery [27], así como las referencias al final del libro. También es útil el material complementario para este capítulo.

1. Identificación y enunciación del problema. Este punto podría parecer muy obvio, pero es común que en la práctica no sea sencillo darse cuenta de que existe un problema que requiere experimentación, y

Tabla 1-1 Pautas generales para diseñar un experimento

1. Identificación y exposición del problema. 2. Elección de los factores, los niveles y los rangos.[a] 3. Selección de la variable de respuesta.[a]	Planeación previa al experimento
4. Elección del diseño experimental.	
5. Realización del experimento.	
6. Análisis estadístico de los datos.	
7. Conclusiones y recomendaciones.	

[a] En la práctica, los pasos 2 y 3 suelen hacerse simultáneamente o en el orden inverso.

tampoco es fácil desarrollar una enunciación clara, con la que todos estén de acuerdo, de este problema. Es necesario desarrollar todas las ideas acerca de los objetivos del experimento. Generalmente, es importante solicitar aportaciones de todas las áreas involucradas: ingeniería, aseguramiento de calidad, manufactura, mercadotecnia, administración, el cliente y el personal de operación (el cual por lo general conoce a fondo el proceso y al que con demasiada frecuencia se ignora). Por esta razón, se recomienda un **enfoque de equipo** para diseñar experimentos.

En la mayoría de los casos es conveniente hacer una lista de los problemas o las preguntas específicas que van a abordarse en el experimento. Una enunciación clara del problema contribuye sustancialmente a menudo para alcanzar una mejor comprensión de los fenómenos bajo estudio y la solución final del problema. También es importante tener presente el objetivo global; por ejemplo, ¿se trata de un proceso o sistema nuevo (en cuyo caso el objetivo inicial posiblemente será la caracterización o tamizado de los factores) o se trata de un sistema maduro que se conoce con profundidad razonable y que se ha caracterizado con anterioridad (en cuyo caso el objetivo puede ser la optimización)? En un experimento puede haber muchos objetivos posibles, incluyendo la confirmación (¿el sistema se comporta de la misma manera ahora que en el pasado?), el descubrimiento (¿qué ocurre si se exploran nuevos materiales, variables, condiciones de operación, etc.?) y la estabilidad (¿bajo qué condiciones las variables de respuesta de interés sufren una degradación seria?). Obviamente, las cuestiones específicas que habrán de abordarse en el experimento se relacionan de manera directa con los objetivos globales. Con frecuencia en esta etapa de la formulación del problema muchos ingenieros y científicos se percatan de que no es posible que un experimento **comprensivo** extenso responda las cuestiones clave y de que un **enfoque secuencial** en el que se utilice una serie de experimentos más pequeños es una estrategia más adecuada.

2. Elección de los factores, los niveles y los rangos. (Como se indica en la tabla 1-1, los pasos 2 y 3 muchas veces se hacen simultáneamente o en orden inverso.) Cuando se consideran los factores que pueden influir en el desempeño de un proceso o sistema, el experimentador suele descubrir que estos factores pueden clasificarse como **factores potenciales del diseño** o bien como factores perturbadores. Los factores potenciales del diseño son aquellos que el experimentador posiblemente quiera hacer variar en el experimento. Es frecuente encontrar que hay muchos factores potenciales del diseño, por lo que es conveniente contar con alguna clasificación adicional de los mismos. Algunas clasificaciones útiles son **factores del diseño, factores que se mantienen constantes** y **factores a los que se permite variar.** Los factores del diseño son los que se seleccionan realmente para estudiarlos en el experimento. Los factores que se mantienen constantes son variables que pueden tener cierto efecto sobre la respuesta, pero que para los fines del experimento en curso no son de interés, por lo que se mantendrán fijos en un nivel específico. Por ejemplo, en un experimento de grabado químico en la industria de los semiconductores puede haber un efecto, que es único, de la herramienta específica para el grabado químico con plasma que se utiliza en el experimento. Sin embargo, sería muy difícil variar este factor en un experimento, por lo que el experimentador puede decidir llevar a cabo todas las corridas experimentales en un grabador químico particular (idealmente "típico"). De este modo, este factor se mantiene constante. Como un ejemplo de factores

a los que se permite variar, las unidades experimentales o los "materiales" a los que se aplican los factores del diseño no son homogéneos por lo general, no obstante lo cual con frecuencia se ignora esta variabilidad de una unidad a otra y se confía en la aleatorización para compensar cualquier efecto del material o la unidad experimental. Muchas veces se trabajará con el supuesto de que los efectos de los factores que se mantienen constantes y de los factores a los que se permite variar son relativamente pequeños.

Por otra parte, los factores perturbadores pueden tener efectos considerables que deben tomarse en consideración, a pesar de que no haya interés en ellos en el contexto del experimento en curso. Los factores perturbadores suelen clasificarse como **factores controlables, no controlables** o **de ruido.** Un factor perturbador controlable es aquel cuyos niveles pueden ser ajustados por el experimentador. Por ejemplo, el experimentador puede seleccionar lotes diferentes de materia prima o diversos días de la semana para conducir el experimento. La estructura básica de la formación de bloques, comentada en la sección anterior, suele ser útil para trabajar con factores perturbadores controlables. Si un factor perturbador no es controlable en el experimento, pero puede medirse, muchas veces puede usarse el procedimiento de análisis denominado **análisis de covarianza** para compensar este efecto. Por ejemplo, la humedad relativa en el medio ambiente del proceso puede afectar el desempeño del proceso, y si la humedad no puede controlarse, probablemente podrá medirse y tratarse como una covariable. Cuando un factor que varía de manera natural y no controlable en el proceso puede controlarse para los fines de un experimento, con frecuencia se le llama factor de ruido. En tales situaciones, es común que el objetivo sea encontrar los ajustes de los factores controlables del diseño que minimicen la variabilidad transmitida por los factores de ruido. En ocasiones a esto se le llama el estudio de robustez del proceso o el problema de robustez del diseño. La formación de bloques, el análisis de covarianza y los estudios de robustez del proceso se comentan más adelante.

Una vez que el experimentador ha seleccionado los factores del diseño, debe elegir los rangos en los que hará variar estos factores, así como los niveles específicos con los que se realizarán las corridas. También deberá pensarse cómo van a controlarse estos factores en los valores deseados y cómo van a medirse. Por ejemplo, en el experimento de la soldadura líquida, el ingeniero ha definido 12 variables que pueden afectar la ocurrencia de defectos de soldadura. El ingeniero también tendrá que tomar una decisión en cuanto a la región de interés para cada variable (es decir, el rango en el que se hará variar cada factor) y en cuanto al número de niveles de cada variable que usará. Para ello se requiere del **conocimiento del proceso.** Este conocimiento del proceso suele ser una combinación de experiencia práctica y conocimientos teóricos. Es importante investigar todos los factores que pueden ser de importancia y no dejarse influir demasiado por la experiencia pasada, en particular cuando uno se encuentra en las fases iniciales de la experimentación o cuando el proceso no está del todo maduro.

Cuando el objetivo del experimento es el tamizado de los factores o caracterización del proceso, por lo general es mejor mantener reducido el número de niveles de los factores. En general, dos niveles funcionan bastante bien en los estudios de tamizado de factores. Elegir la región de interés también es importante. En el tamizado de factores, la región de interés deberá ser relativamente grande; es decir, el rango en el que se hacen variar los factores deberá ser amplio. Conforme se sepa más acerca de las variables que son importantes y de los niveles que producen los mejores resultados, la región de interés se hará por lo general más estrecha.

3. Selección de la variable de respuesta. Para seleccionar la variable de respuesta, el experimentador deberá tener la certeza de que esta variable proporciona en realidad información útil acerca del proceso bajo estudio. En la mayoría de los casos, el promedio o la desviación estándar (o ambos) de la característica medida será la variable de respuesta. No son la excepción las respuestas múltiples. La eficiencia de los instrumentos de medición (o error de medición) también es un factor importante. Si la eficiencia de los instrumentos de medición es inadecuada, el experimentador sólo detectará los efectos relativamente grandes de los factores o quizá sean necesarias réplicas adicionales. En algunas situaciones en que la eficiencia de los instrumentos de medición es pobre, el experimentador puede decidir medir varias veces cada unidad

experimental y usar el promedio de las mediciones repetidas como respuesta observada. Suele ser de importancia determinante identificar los aspectos relacionados con la definición de las respuestas de interés y cómo van a medirse antes de llevar a cabo el experimento. En ocasiones se emplean experimentos diseñados para estudiar y mejorar el desempeño de los sistemas de medición. Para un ejemplo, ver el capítulo 12.

Se reitera lo crucial que es exponer todos los puntos de vista y la información del proceso en los pasos 1 al 3 anteriores. Se hace referencia a esto como **planeación previa al experimento.** Coleman y Montgomery [27] proporcionan hojas de trabajo que pueden ser útiles en la planeación previa al experimento. Véase también la información complementaria del texto para más detalles y un ejemplo del uso de estas hojas de trabajo. En muchas situaciones, no es posible que una sola persona posea todos los conocimientos requeridos para hacer esto adecuadamente. Por lo tanto, se hace una amplia recomendación para el trabajo en equipo durante la planeación del experimento. La mayor parte del éxito gravitará en torno a qué tan bien se haya hecho la planeación previa del experimento.

4. Elección del diseño experimental. Si las actividades de planeación previas al experimento se realizan como es debido, este paso es relativamente sencillo. La elección del diseño implica la consideración del tamaño de la muestra (número de réplicas), la selección de un orden de corridas adecuado para los ensayos experimentales y la determinación de si entran en juego o no la formación de bloques u otras restricciones sobre la aleatorización. En este libro se revisan algunos de los tipos más importantes de diseños experimentales, y puede usarse en última instancia como un catálogo para seleccionar el diseño experimental apropiado para una amplia variedad de problemas.

Existen también varios paquetes interactivos de software de estadística que soportan esta fase del diseño experimental. El experimentador puede introducir la información del número de factores, los niveles y los rangos, y estos programas presentarán a la consideración del experimentador una selección de diseños o recomendarán un diseño particular. (Nosotros preferimos ver varias alternativas en lugar de confiar en la recomendación de la computadora en la mayoría de los casos.) Estos programas proporcionan también por lo general una hoja de trabajo (con el orden aleatorizado de las corridas) que se usará en la conducción del experimento.

Al seleccionar el diseño, es importante tener en mente los objetivos experimentales. En muchos experimentos de ingeniería se sabe de antemano que algunos de los niveles de los factores producirán valores diferentes de la respuesta. En consecuencia, el interés se centra en identificar *qué* factores causan esta diferencia y en estimar la *magnitud* del cambio de la respuesta. En otras situaciones podría haber más interés en verificar la uniformidad. Por ejemplo, pueden compararse dos condiciones de producción A y B, donde A es el estándar y B es una alternativa con una eficiencia de costos mayor. El experimentador estará interesado entonces en demostrar que, por ejemplo, no hay ninguna diferencia en el rendimiento entre las dos condiciones.

5. Realización del experimento. Cuando se lleva a cabo el experimento es vital monitorear con atención el proceso a fin de asegurarse de que todo se esté haciendo conforme a la planeación. Los errores en el procedimiento experimental en esta etapa destruirán por lo general la validez experimental. Poner en un primer plano la planeación es crucial para el éxito. Es fácil subestimar los aspectos de logística y planeación cuando se corre un experimento diseñado en un ambiente complejo de manufactura o de investigación y desarrollo.

Coleman y Montgomery [27] sugieren que antes de llevar a cabo el experimento, es conveniente en muchas ocasiones realizar algunas corridas piloto o de prueba. Estas corridas proporcionan información acerca de la consistencia del material experimental, una comprobación del sistema de medición, una idea aproximada del error experimental y la oportunidad de poner en práctica la técnica experimental global. Esto ofrece también una oportunidad para revisar, de ser necesario, las decisiones tomadas en los pasos 1 al 4.

6. Análisis estadístico de los datos. Deberán usarse métodos estadísticos para analizar los datos a fin de que los resultados y las conclusiones sean objetivos y no de carácter apreciativo. Si el experimento se ha diseñado correctamente y si se ha llevado a cabo de acuerdo con el diseño, los métodos estadísticos necesarios no deben ser complicados. Existen varios paquetes de software excelentes diseñados para auxiliar en el análisis de datos, y muchos de los programas usados en el paso 4 para seleccionar el diseño cuentan con una interfase directa para el análisis estadístico. Con frecuencia se encuentra que los métodos gráficos simples desempeñan un papel importante en el análisis e interpretación de datos. Debido a que muchas de las preguntas que el experimentador quiere responder pueden insertarse en el marco de la prueba de hipótesis, los procedimientos para probar hipótesis y estimar intervalos de confianza son muy útiles en el análisis de datos de un experimento diseñado. Muchas veces es muy útil también presentar los resultados de varios experimentos en términos de un **modelo empírico**, es decir, mediante una ecuación derivada de los datos que expresa la relación entre la respuesta y los factores importantes del diseño. El análisis residual y la verificación de la adecuación del modelo son también técnicas de análisis importantes. Más adelante se revisarán en detalle estos temas.

Recuerde que los métodos estadísticos no pueden demostrar que un factor (o factores) posee un efecto particular, sólo proporcionan pautas generales en cuanto a la confiabilidad y la validez de los resultados. Aplicados en forma correcta, los métodos estadísticos no permiten la demostración experimental de nada, pero sí sirven para medir el error posible en una conclusión o asignar un nivel de confianza a un enunciado. La ventaja principal de los métodos estadísticos es que agregan objetividad al proceso de toma de decisiones. Las técnicas estadísticas, aunadas a una buena ingeniería o conocimiento del proceso y el sentido común, llevarán por lo general a conclusiones sólidas.

7. Conclusiones y recomendaciones. Una vez que se han analizado los datos, el experimentador debe sacar conclusiones *prácticas* acerca de los resultados y recomendar un curso de acción. Los métodos gráficos suelen ser útiles en esta etapa, en particular para presentar los resultados. También deberán realizarse **corridas de seguimiento** o **pruebas de confirmación** para validar las conclusiones del experimento.

A lo largo del proceso completo es importante tener presente que la experimentación es una parte esencial del proceso de aprendizaje, en la que se formulan hipótesis tentativas acerca de un sistema, se realizan experimentos para investigar estas hipótesis y se formulan nuevas hipótesis con base en los resultados, y así sucesivamente. Esto sugiere que la experimentación es **iterativa**. Por lo general es un gran error diseñar un solo experimento comprensivo y extenso al principio de un estudio. Un experimento exitoso requiere conocer los factores importantes, los rangos en los que deberán hacerse variar estos factores, el número apropiado de niveles que deberán usarse y las unidades de medición apropiadas para estas variables. En general, no se conocen las respuestas precisas de estas cuestiones, pero se aprende acerca de ellas sobre la marcha. A medida que avanza un programa experimental, es común abandonar algunas variables de entrada e incorporar otras, modificar la región de exploración de algunos factores o incorporar nuevas variables de respuesta. Por consiguiente, generalmente la experimentación se hace en forma secuencial y, como regla general, no deberá invertirse más de 25% de los recursos disponibles en el primer experimento. Con esto se asegurará que se contará con los recursos suficientes para realizar las corridas de confirmación y que se alcanzará en última instancia el objetivo final del experimento.

1-5 BREVE HISTORIA DEL DISEÑO ESTADÍSTICO

Ha habido cuatro eras del desarrollo moderno del diseño experimental estadístico. La era agrícola fue encabezada por el trabajo pionero de Sir Ronald A. Fisher en los años 1920 y principios de la década de 1930. En este periodo, Fisher fue el responsable de las estadísticas y el análisis de datos en la Estación

Agrícola Experimental de Rothamsted en las cercanías de Londres, Inglaterra. Fisher se percató de que las fallas en la forma en que se llevaba a cabo el experimento que generaba los datos obstaculizaban con frecuencia el análisis de los datos de los sistemas (en este caso sistemas agrícolas). Mediante la interacción con múltiples científicos e investigadores de diversos campos, Fisher desarrolló las ideas que llevaron a los tres principios básicos del diseño experimental que se revisan en la sección 1-3: la aleatorización, la realización de réplicas y la formación de bloques. Fisher incorporó de manera sistemática el pensamiento y los principios estadísticos en el diseño de las investigaciones experimentales, incluyendo el concepto de diseño factorial y el análisis de varianza. Sus libros [44a, b] tuvieron profunda influencia en el uso de la estadística, particularmente en la agricultura y las ciencias biológicas relacionadas. Para una excelente biografía de Fisher, ver Box [21].

Si bien es cierto que la aplicación del diseño estadístico en ambientes industriales se inició en la década de 1930, el catalizador de la segunda era, o era industrial, fue el desarrollo de la metodología de superficies de respuesta (MSR) por parte de Box y Wilson [20]. Estos autores se percataron y explotaron el hecho de que muchos experimentos industriales son fundamentalmente diferentes de sus contrapartes agrícolas en dos sentidos: 1) la variable de respuesta puede observarse por lo general (casi) de inmediato, y 2) el experimentador puede obtener pronto información crucial de un pequeño grupo de corridas que puede usarse para planear el siguiente experimento. Box [12f] denomina *inmediatez* y *secuencialidad* a estas dos características de los experimentos industriales. En los 30 años siguientes, la MSR y otras técnicas de diseño se generalizaron en las industrias química y de proceso, sobre todo en el trabajo de investigación y desarrollo. George Box fue el líder intelectual de este movimiento. Sin embargo, la aplicación del diseño estadístico a nivel de plantas o procesos de manufactura todavía no estaba muy generalizada. Algunas de las razones de ello incluyen la capacitación inadecuada de ingenieros y otros especialistas en procesos en los conceptos y los métodos estadísticos básicos, así como la falta de recursos de computación y software de estadística que fueran fáciles de usar para apoyar la aplicación de experimentos diseñados estadísticamente.

El interés creciente de la industria occidental en el mejoramiento de calidad que empezó a fines de la década de 1970 anunció la tercera era del diseño estadístico. El trabajo de Genichi Taguchi (Taguchi y Wu [109], Kackar [62] y Taguchi [108a, b]) tuvo un impacto significativo en el aumento del interés y el uso de los experimentos diseñados. Taguchi propugnaba por el uso de experimentos diseñados para lo que denominó el diseño paramétrico robusto, es decir,

1. Hacer procesos insensibles a los factores ambientales o de otra índole que son difíciles de controlar.
2. Fabricar productos insensibles a la variación transmitida por los componentes.
3. Encontrar los niveles de las variables del proceso que obliguen a la media a un valor deseado mientras que al mismo tiempo se reduzca la variabilidad en torno a este valor.

Taguchi propuso diseños factoriales altamente fraccionados y otros arreglos ortogonales junto con algunos métodos estadísticos nuevos para resolver estos problemas. La metodología resultante generó muchas discusiones y controversias. Parte de la controversia surgió porque en Occidente la metodología de Taguchi fue defendida al principio (y sobre todo) por empresarios, y no se había hecho la revisión escrutadora adecuada de la ciencia estadística fundamental. Para fines de la década de 1980, los resultados de esta revisión indicaron que aun cuando los conceptos y los objetivos enfocados en la ingeniería de Taguchi tenían bases sólidas, existían problemas sustanciales con su estrategia experimental y sus métodos para el análisis de los datos. Para detalles específicos de estas cuestiones, ver Box [12d], Box, Bisgaard y Fung [14], Hunter [59a, b], Myers y Montgomery [85a] y Pignatiello y Ramberg [94]. Gran parte de estas preo-

cupaciones se resumen también en el amplio panel de discusión del número de mayo de 1992 de *Technometrics* (ver Nair, *et al.* [86]).

Hubo al menos tres resultados positivos de la controversia desatada por Taguchi. Primero, el uso de los experimentos diseñados se hizo más generalizado en las industrias con piezas discretas, incluyendo la industria de manufacturas automotrices y aeroespaciales, de electrónica y semiconductores, y muchas otras, que anteriormente hacían poco uso de esta técnica. Segundo, se inició la cuarta era del diseño estadístico. Esta era ha incluido un renovado interés general tanto por parte de investigadores como de profesionales en ejercicio en el diseño estadístico y el desarrollo de varios enfoques nuevos y útiles para los problemas experimentales en el mundo industrial, incluyendo alternativas a los métodos técnicos de Taguchi que permiten que sus conceptos de ingeniería se lleven a la práctica de manera eficaz y eficiente. Algunas de estas alternativas se revisarán e ilustrarán en capítulos subsecuentes, en particular en el capítulo 11. Tercero, la educación formal en diseño experimental estadístico se está haciendo parte de los programas de ingeniería en las universidades, tanto a nivel de licenciatura como de posgrado. La integración exitosa de una buena práctica del diseño experimental en la ingeniería y las ciencias es un factor clave en la competitividad industrial futura.

1-6 RESUMEN: USO DE TÉCNICAS ESTADÍSTICAS EN LA EXPERIMENTACIÓN

Gran parte de la investigación en la ingeniería, las ciencias y la industria es empírica y hace un uso extensivo de la experimentación. Los métodos estadísticos pueden incrementar en gran medida la eficiencia de estos experimentos y con frecuencia pueden fortalecer las conclusiones así obtenidas. El uso correcto de las técnicas estadísticas en la experimentación requiere que el experimentador tenga presentes los puntos siguientes:

1. ***Uso de conocimientos no estadísticos del problema.*** Los experimentadores suelen poseer amplios conocimientos de sus respectivos campos. Por ejemplo, un ingeniero civil que trabaja en un problema de hidrología cuenta de manera típica con considerable experiencia práctica y capacitación académica formal en esta área. En algunos campos existe un cuerpo enorme de teoría física en el cual indagar para explicar las relaciones entre los factores y las respuestas. Este tipo de conocimientos no estadísticos es invaluable para elegir los factores, determinar los niveles de los factores, decidir cuántas réplicas correr, interpretar los resultados del análisis, etc. El uso de la estadística no es sustituto de la reflexión sobre el problema.

2. ***Mantener el diseño y el análisis tan simple como sea posible.*** Es necesario no exagerar en el uso de técnicas estadísticas complejas y sofisticadas. Los métodos de diseño y análisis relativamente simples son siempre los mejores. En este punto cabe hacer hincapié nuevamente en el paso 4 del procedimiento recomendado en la sección 1-4. Si un diseño se hace de manera cuidadosa y correcta, el análisis casi siempre será relativamente directo. Sin embargo, si el diseño se estropea grandemente por ineptitud, no es posible que incluso la estadística más compleja y elegante salve la situación.

3. ***Tener presente la diferencia entre significación práctica y significación estadística.*** Debido justamente a que dos condiciones experimentales producen respuestas medias que son estadísticamente diferentes, no existe ninguna seguridad de que esta diferencia sea de la magnitud suficiente como para tener algún valor práctico. Por ejemplo, un ingeniero puede determinar que una modificación en el sistema de inyección de combustible de un automóvil puede producir un mejoramiento promedio real en el rendimiento del combustible de 0.1 mi/gal. Éste es un resultado estadísticamente significativo. Sin embargo, si

el costo de la modificación es de $1000, la diferencia de 0.1 mi/gal probablemente será muy pequeña para poseer algún valor práctico.

4. Los experimentos son generalmente iterativos. Recuerde que en la mayoría de las situaciones no es conveniente diseñar un experimento demasiado comprensivo al principio de un estudio. Un diseño exitoso requiere conocer los factores importantes, los rangos en los que estos factores se harán variar, el número apropiado de niveles para cada factor y los métodos y las unidades de medición adecuados para cada factor y respuesta. En general, ningún experimentador está en posición de responder estas cuestiones al principio del experimento, sino que las respuestas aparecen sobre la marcha. Esto habla en favor del enfoque *iterativo* o *secuencial* analizado anteriormente. Desde luego, hay situaciones en las que un experimento comprensivo es totalmente apropiado pero, como regla general, la mayoría de los experimentos deberán ser iterativos. Por consiguiente, no deberá invertirse más de 25% de los recursos para la experimentación (corridas, presupuesto, tiempo, etc.) en el experimento inicial. Con frecuencia estos esfuerzos iniciales constituyen sólo experiencias de aprendizaje, y es necesario contar con recursos suficientes para alcanzar los objetivos finales del experimento.

2 Experimentos comparativos simples

En este capítulo se examinan los experimentos para comparar dos **condiciones** (llamadas en ocasiones **tratamientos**), a las cuales es común denominar **experimentos comparativos simples**. Se empieza con el ejemplo de un experimento que se realiza para determinar si dos formulaciones diferentes de un producto producen resultados equivalentes. El estudio lleva a revisar varios conceptos básicos de la estadística, como variables aleatorias, distribuciones de probabilidad, muestras aleatorias, distribuciones de muestreo y pruebas de hipótesis.

2-1 INTRODUCCIÓN

La fuerza de la tensión de adhesión del mortero de cemento portland es una característica importante del producto. Un ingeniero está interesado en comparar la fuerza de una formulación modificada en la que se han agregado emulsiones de látex de polímeros durante el mezclado, con la fuerza del mortero sin modificar. El experimentador ha reunido 10 observaciones de la fuerza de la formulación modificada y otras 10 observaciones de la formulación sin modificar. Los datos se muestran en la tabla 2-1. Podría hacerse referencia a las dos formulaciones diferentes como dos **tratamientos** o como dos **niveles** del **factor** formulaciones.

En la figura 2-1 se grafican los datos de este experimento. A esta representación se le llama **diagrama de puntos**. Del examen visual de estos datos se obtiene la impresión inmediata de que la fuerza del mortero sin modificar es mayor que la fuerza del mortero modificado. Esta impresión se confirma al comparar las fuerzas de la tensión de adhesión *promedio*, $\bar{y}_1 = 16.76$ kgf/cm^2 para el mortero modificado y $\bar{y}_2 = 17.92$ kgf/cm^2 para el mortero sin modificar. Las fuerzas de la tensión de adhesión promedio de estas dos muestras difieren en lo que parece ser una cantidad no trivial. Sin embargo, no es evidente que esta diferencia sea de la magnitud suficiente para implicar que las dos formulaciones *son* en realidad diferentes. Quizás esta diferencia observada en las fuerzas promedio sea el resultado de fluctuaciones del muestreo y las dos formulaciones sean idénticas en realidad. Posiblemente otras dos muestras produzcan el resultado contrario, con la fuerza del mortero modificado excediendo la de la formulación sin modificar.

Puede usarse una técnica de la inferencia estadística llamada **prueba de hipótesis** (algunos autores prefieren el término **prueba de significación**) para auxiliar al experimentador en la comparación de estas

Tabla 2-1 Datos de la fuerza de la tensión de adhesión del experimento de la formulación del cemento portland

j	Mortero modificado y_{1j}	Mortero sin modificar y_{2j}
1	16.85	17.50
2	16.40	17.63
3	17.21	18.25
4	16.35	18.00
5	16.52	17.86
6	17.04	17.75
7	16.96	18.22
8	17.15	17.90
9	16.59	17.96
10	16.57	18.15

dos formulaciones. La prueba de hipótesis permite que la comparación de las dos formulaciones se haga en términos *objetivos*, con el conocimiento de los riesgos asociados si se llega a una conclusión equivocada. Antes de presentar los procedimientos de la prueba de hipótesis en experimentos comparativos simples, se hará una breve revisión de algunos conceptos elementales de la estadística.

2-2 CONCEPTOS ESTADÍSTICOS BÁSICOS

A cada una de las observaciones del experimento del cemento portland citado anteriormente se le llamaría una **corrida**. Observe que las corridas individuales difieren, por lo que existen fluctuaciones, o **ruido**, en los resultados. Es común llamar a este ruido el **error experimental** o simplemente el **error**. Se trata de un **error estadístico**, lo cual significa que se origina por la variación que no está bajo control y que generalmente es inevitable. La presencia del error o ruido implica que la variable de respuesta, la fuerza de la tensión de adhesión, es una **variable aleatoria**. Una variable aleatoria puede ser **discreta** o **continua**. Si el conjunto de todos los valores posibles de la variable aleatoria es finito o contablemente infinito, entonces la variable aleatoria es discreta, mientras que si el conjunto de todos los valores posibles de la variable aleatoria es un intervalo, entonces la variable aleatoria es continua.

Descripción gráfica de la variabilidad

Es frecuente usar métodos gráficos simples como ayuda para analizar los datos de un experimento. El **diagrama de puntos**, ilustrado en la figura 2-1, es un recurso muy útil para representar un cuerpo reducido de datos (digamos hasta unas 20 observaciones). El diagrama de puntos le permite al experimentador ver de inmediato la **localización** o **tendencia central** de las observaciones y su **dispersión**. Por ejemplo, en el experimento de la fuerza de tensión de adhesión del cemento portland, el diagrama de puntos revela

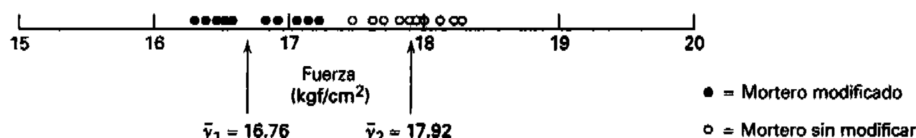


Figura 2-1 Diagrama de puntos de los datos de la fuerza de la tensión de adhesión de la tabla 2-1.

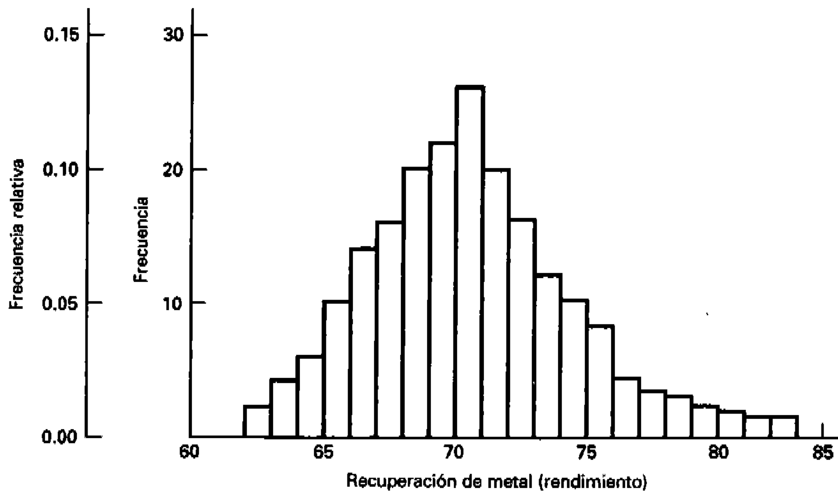


Figura 2-2 Histograma de 200 observaciones de la recuperación de metal (rendimiento) en un proceso de fundición.

que probablemente las dos formulaciones difieran en la fuerza promedio, pero que ambas producen aproximadamente la misma variación en la fuerza.

Cuando los datos son muy numerosos, es difícil distinguir las observaciones graficadas en un diagrama de puntos, y en tal caso sería preferible un histograma. En la figura 2-2 se presenta el histograma de 200 observaciones de la recuperación de metal (o rendimiento) en un proceso de fundición. El histograma muestra la tendencia central, la dispersión y la forma general de la distribución de los datos. Recuerde que un histograma se construye dividiendo el eje horizontal en intervalos (generalmente de longitud igual) y trazando un rectángulo sobre el intervalo j-ésimo con el área del rectángulo proporcional a n_j, el número de observaciones incluidas en ese intervalo.

El **diagrama de caja** (o **diagrama de caja y bigotes**) es una manera muy útil de representar gráficamente los datos. En un diagrama de caja se muestra el mínimo, el máximo, los cuartiles inferior y superior (el percentil 25 y el percentil 75, respectivamente) y la mediana (el percentil 50) en una caja rectangular alineada horizontal o verticalmente. La caja se extiende del cuartil inferior al cuartil superior y se traza una línea por la mediana que atraviesa la caja. Se trazan dos líneas (o bigotes) que se extienden de los extremos de la caja hasta (de manera típica) los valores mínimo y máximo. (Existen diversas variantes de los diagramas de caja que tienen reglas diferentes para denotar los puntos muestrales extremos. Ver Montgomery y Runger [83d] para más detalles.)

En la figura 2-3 se muestran los diagramas de caja de las dos muestras de la fuerza de la tensión de adhesión en el experimento del mortero de cemento portland. En esta representación se revela con toda claridad la diferencia en la fuerza promedio entre las dos formulaciones. Indica asimismo que ambas formulaciones producen distribuciones de la fuerza razonablemente simétricas con una variabilidad o dispersión similar.

Los diagramas de puntos, los histogramas y los diagramas de caja son útiles para resumir la información de una **muestra** de datos. Para describir con mayor detalle las observaciones que podrían presentarse en una muestra, se usa el concepto de distribución de probabilidad.

Distribuciones de probabilidad

La estructura de la probabilidad de una variable aleatoria, por ejemplo y, se describe mediante su **distribución de probabilidad**. Cuando y es discreta, es común hacer referencia a su distribución de probabili-

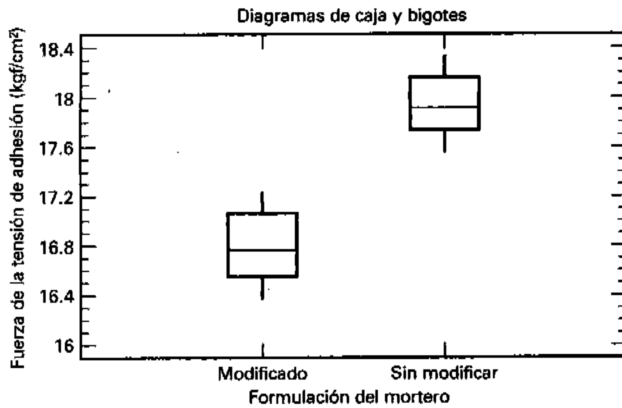
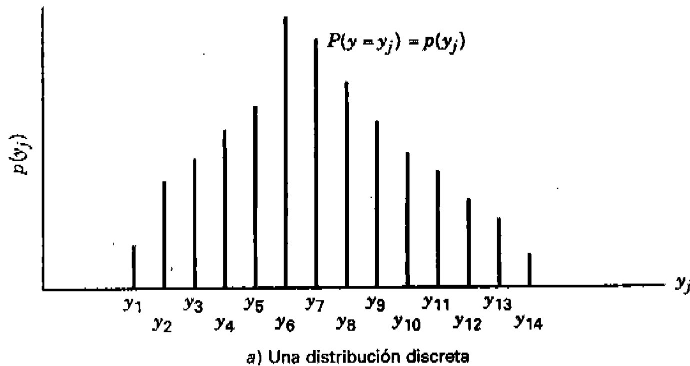


Figura 2-3 Diagramas de caja del experimento de la fuerza de la tensión de adhesión del cemento portland.

dad, por ejemplo $p(y)$, como la función de probabilidad de y. Cuando y es continua, es común hacer referencia a su distribución de probabilidad, por ejemplo $f(y)$, como la función de densidad de probabilidad de y.

En la figura 2-4 se ilustran dos distribuciones de probabilidad hipotéticas, una discreta y la otra continua. Observe que en la distribución de probabilidad discreta es la altura de la función $p(y_j)$ la que representa la probabilidad, mientras que en el caso continuo, es el área bajo la curva $f(y)$ asociada con un



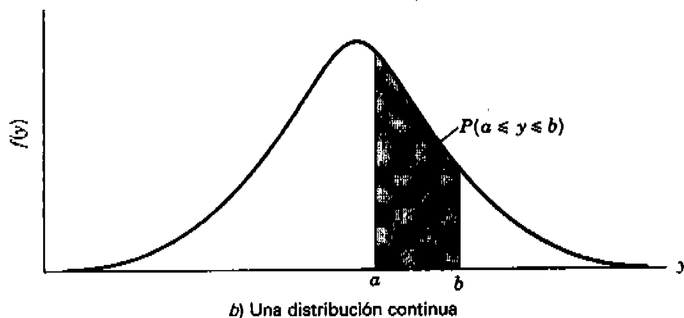


Figura 2-4 Distribuciones de probabilidad discreta y continua.

intervalo dado la que representa la probabilidad. Un resumen cuantitativo de las propiedades de las distribuciones de probabilidad sería el siguiente:

y discreta:

$$0 \le p(y_j) \le 1 \qquad \text{todos los valores de } y_j$$
$$P(y = y_j) = p(y_j) \qquad \text{todos los valores de } y_j$$
$$\sum_{\substack{\text{todos los} \\ \text{valores de } y_j}} p(y_j) = 1$$

y continua:

$$0 \le f(y)$$

$$P(a \le y \le b) = \int_a^b f(y)\,dy$$

$$\int_{-\infty}^{\infty} f(y)\,dy = 1$$

Media, varianza y valores esperados

La **media,** μ, de una distribución de probabilidad es una medida de su tendencia central o localización. Matemáticamente, la media se define como

$$\mu = \begin{cases} \int_{-\infty}^{\infty} y f(y)\,dy & y \text{ continua} \\ \sum_{\text{toda } y} y p(y) & y \text{ discreta} \end{cases} \tag{2-1}$$

La media también puede expresarse en términos del **valor esperado** o valor promedio a la larga de la variable aleatoria y como

$$\mu = E(y) = \begin{cases} \int_{-\infty}^{\infty} y f(y)\,dy & y \text{ continua} \\ \sum_{\text{toda } y} y p(y) & y \text{ discreta} \end{cases} \tag{2-2}$$

donde E denota el **operador del valor esperado.**

La variabilidad o dispersión de una distribución de probabilidad puede medirse con la **varianza,** la cual se define como

$$\sigma^2 = \begin{cases} \int_{-\infty}^{\infty} (y-\mu)^2 f(y)\,dy & y \text{ continua} \\ \sum_{\text{toda } y} (y-\mu)^2 p(y) & y \text{ discreta} \end{cases} \tag{2-3}$$

Observe que la varianza puede expresarse exclusivamente en términos del valor esperado debido a que

$$\sigma^2 = E[(y-\mu)^2] \tag{2-4}$$

Por último, el uso de la varianza es tan frecuente que resulta conveniente definir un **operador de la varianza** V tal que

$$V(y) \equiv E[(y-\mu)^2] = \sigma^2 \tag{2-5}$$

Los conceptos de valor esperado y varianza se usan constantemente a lo largo de este libro, y puede ser útil revisar varios resultados elementales relacionados con estos operadores. Si y es una variable aleatoria con media μ y varianza σ^2 y c es una constante, entonces

1. $E(c) = c$
2. $E(y) = \mu$
3. $E(cy) = cE(y) = c\mu$
4. $V(c) = 0$
5. $V(y) = \sigma^2$
6. $V(cy) = c^2V(y) = c^2\sigma^2$

Si hay dos variables aleatorias, por ejemplo, y_1 con $E(y_1) = \mu_1$ y $V(y_1) = \sigma_1^2$ y y_2 con $E(y_2) = \mu_2$ y $V(y_2) = \sigma_2^2$, se tiene

7. $E(y_1 + y_2) = E(y_1) + E(y_2) = \mu_1 + \mu_2$

Es posible demostrar que

8. $V(y_1 + y_2) = V(y_1) + V(y_2) + 2\ \text{Cov}(y_1, y_2)$

donde

$$\text{Cov}(y_1, y_2) = E[(y_1 - \mu_1)(y_2 - \mu_2)] \tag{2-6}$$

es la **covarianza** de las variables aleatorias y_1 y y_2. La covarianza es una medida de la asociación lineal entre y_1 y y_2. Más específicamente, puede demostrarse que si y_1 y y_2 son independientes,[1] entonces $\text{Cov}(y_1, y_2) = 0$. También puede demostrarse que

9. $V(y_1 - y_2) = V(y_1) + V(y_2) - 2\ \text{Cov}(y_1, y_2)$

Si y_1 y y_2 son **independientes**, se tiene

10. $V(y_1 \pm y_2) = V(y_1) + V(y_2) = \sigma_1^2 + \sigma_2^2$

y

11. $E(y_1 \cdot y_2) = E(y_1) \cdot E(y_2) = \mu_1 \cdot \mu_2$

Sin embargo, observe que, en general,

12. $E\left(\dfrac{y_1}{y_2}\right) \neq \dfrac{E(y_1)}{E(y_2)}$

sin importar si y_1 y y_2 son independientes o no.

2-3 MUESTREO Y DISTRIBUCIONES DE MUESTREO

Muestras aleatorias, media muestral y varianza muestral

El objetivo de la inferencia estadística es sacar conclusiones acerca de una población utilizando una muestra de la misma. La mayoría de los métodos que se estudiarán aquí incluyen el supuesto de que se

[1] Observe que el recíproco no es necesariamente verdadero; es decir, puede tenerse $\text{Cov}(y_1, y_2) = 0$ y no obstante esto no implica que las variables sean independientes. Para un ejemplo, ver Hines y Montgomery ([55] pp. 128-129).

usan muestras aleatorias. Es decir, si la población contiene N elementos y va a seleccionarse una muestra de n de ellos, y si cada una de las $N!/(N-n)!n!$ muestras posibles tiene una probabilidad igual de ser escogida, entonces al procedimiento empleado se le llama **muestreo aleatorio**. En la práctica, en ocasiones es difícil obtener muestras aleatorias, para lo cual pueden ser útiles las tablas de números aleatorios, como la tabla XI del apéndice.

En la inferencia estadística se utilizan profusamente cantidades calculadas a partir de las observaciones de la muestra. Un **estadístico** se define como cualquier función de las observaciones de una muestra que no contiene parámetros desconocidos. Por ejemplo, suponga que $y_1, y_2, ..., y_n$ representa una muestra. Entonces la **media muestral**

$$\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n} \tag{2-7}$$

y la **varianza muestral**

$$S^2 = \frac{\sum_{i=1}^{n} (y_i - \bar{y})^2}{n-1} \tag{2-8}$$

son estadísticos. Estas cantidades son medidas de la tendencia central y la dispersión de la muestra, respectivamente. En ocasiones se usa $S = \sqrt{S^2}$, llamada la **desviación estándar muestral**, como medida de dispersión. Los ingenieros suelen preferir el uso de la desviación estándar para medir la dispersión debido a que se expresa en las mismas unidades que la variable de interés y.

Propiedades de la media y la varianza muestrales

La media muestral $\bar{y}$ es un estimador puntual de la media poblacional μ, y la varianza muestral S^2 es un estimador puntual de la varianza poblacional σ^2. En general, un **estimador** de un parámetro desconocido es un estadístico que corresponde con dicho parámetro. Observe que un estimador puntual es una variable aleatoria. Al valor numérico particular de un estimador, calculado a partir de los datos muestrales, se le llama una **estimación**. Por ejemplo, suponga que quiere estimarse la media y la varianza de la resistencia a la ruptura de un tipo particular de fibra textil. Se prueba una muestra aleatoria de $n = 25$ ejemplares de prueba de la fibra y se registra la resistencia de cada uno de ellos. La media y la varianza muestrales se calculan de acuerdo con las ecuaciones 2-7 y 2-8, respectivamente, obteniéndose $\bar{y} = 18.6$ y $S^2 = 1.20$. Por lo tanto, la estimación de μ es $\bar{y} = 18.6$ y la estimación de σ^2 es $S^2 = 1.20$.

Un buen estimador puntual debe tener varias propiedades. Dos de las más importantes son las siguientes:

1. El estimador puntual deberá ser **insesgado**. Es decir, el parámetro que se está estimando deberá ser el promedio o valor esperado a la larga del estimador puntual. Aun cuando la ausencia de sesgo es deseable, esta propiedad por sí sola no siempre hace que un estimador sea adecuado.
2. Un estimador insesgado deberá tener la **varianza mínima**. Esta propiedad establece que el estimador puntual de varianza mínima tiene una varianza que es menor que la varianza de cualquier otro estimador del parámetro en cuestión.

Es sencillo demostrar que $\bar{y}$ y S^2 son estimadores insesgados de μ y σ^2, respectivamente. Considere primero $\bar{y}$. Al utilizar las propiedades del valor esperado, se tiene

$$\begin{aligned}
E(\bar{y}) &= E\left(\frac{\sum_{i=1}^{n} y_i}{n}\right) \\
&= \frac{1}{n} E\left(\sum_{i=1}^{n} y_i\right) \\
&= \frac{1}{n}\sum_{i=1}^{n} E(y_i) \\
&= \frac{1}{n}\sum_{i=1}^{n} \mu \\
&= \mu
\end{aligned}$$

porque el valor esperado de cada observación y_i es μ. Por lo tanto, $\bar{y}$ es un estimador insesgado de μ.

Considere ahora la varianza muestral S^2. Se tiene

$$\begin{aligned}
E(S^2) &= E\left[\frac{\sum_{i=1}^{n} (y_i - \bar{y})^2}{n-1}\right] \\
&= \frac{1}{n-1} E\left[\sum_{i=1}^{n} (y_i - \bar{y})^2\right] \\
&= \frac{1}{n-1} E(SS)
\end{aligned}$$

donde $SS = \Sigma_{i=1}^{n}(y_i - \bar{y})^2$ es la **suma de cuadrados corregida** de las observaciones y_i. Entonces

$$E(SS) = E\left[\sum_{i=1}^{n} (y_i - \bar{y})^2\right] \tag{2-9}$$

$$\begin{aligned}
&= E\left[\sum_{i=1}^{n} y_i^2 - n\bar{y}^2\right] \\
&= \sum_{i=1}^{n} (\mu^2 + \sigma^2) - n(\mu^2 + \sigma^2 / n)
\end{aligned}$$

$$= (n-1)\sigma^2 \tag{2-10}$$

Por lo tanto,

$$\begin{aligned}
E(S^2) &= \frac{1}{n-1} E(SS) \\
&= \sigma^2
\end{aligned}$$

y se observa que S^2 es un estimador insesgado de σ^2.

Grados de libertad

A la cantidad $n - 1$ de la ecuación 2-10 se le llama el **número de grados de libertad** de la suma de cuadrados SS. Se trata de un resultado muy general; es decir, si y es una variable aleatoria con varianza σ^2 y $SS = \Sigma(y_i - \bar{y})^2$ tiene v grados de libertad, entonces

$$E\left(\frac{SS}{v}\right) = \sigma^2 \tag{2-11}$$

El número de grados de libertad de una suma de cuadrados es igual al número de elementos independientes en dicha suma de cuadrados. Por ejemplo, $SS = \Sigma_{i=1}^{n}(y_i - \bar{y})^2$ en la ecuación 2-9 consiste en la suma de los cuadrados de los n elementos $y_1 - \bar{y}$, $y_2 - \bar{y}, \ldots, y_n - \bar{y}$. No todos estos elementos son independientes porque $\Sigma_{i=1}^{n}(y_i - \bar{y}) = 0$; de hecho, sólo $n - 1$ de ellos son independientes, lo cual implica que SS tiene $n - 1$ grados de libertad.

La distribución normal y otras distribuciones de muestreo

En muchas ocasiones la distribución de probabilidad de un estadístico particular puede determinarse si se conoce la distribución de probabilidad de la población de la que se tomó la muestra. A la distribución de probabilidad de un estadístico se le llama la **distribución de muestreo**. A continuación se revisan brevemente varias distribuciones de muestreo útiles.

Una de las distribuciones de muestreo más importantes es la **distribución normal**. Si y es una variable aleatoria normal, la distribución de probabilidad de y es

$$f(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(1/2)[(y-\mu)/\sigma]^2} \qquad -\infty < y < \infty \tag{2-12}$$

donde $-\infty < \mu < \infty$ es la media de la distribución y $\sigma^2 > 0$ es la varianza. En la figura 2-5 se ilustra la distribución normal.

Debido a que las corridas muestrales que difieren como resultado del error experimental a menudo se encuentran descritas adecuadamente en la distribución normal, ésta desempeña un papel fundamental en el análisis de los datos de experimentos diseñados. También es posible definir muchas distribuciones de muestreo importantes en términos de variables aleatorias normales. Con frecuencia se usa la notación $y \sim N(\mu, \sigma^2)$ para denotar que y sigue una distribución normal con media μ y varianza σ^2.

Un caso especial importante de la distribución normal es la **distribución normal estándar**; es decir, $\mu = 0$ y $\sigma^2 = 1$. Se observa que si $y \sim N(\mu, \sigma^2)$, la variable aleatoria

$$z = \frac{y - \mu}{\sigma} \tag{2-13}$$

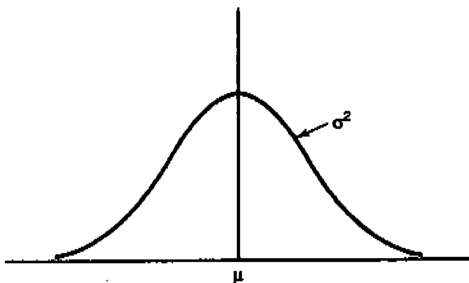


Figura 2-5 La distribución normal.

sigue la distribución normal estándar, denotada $z \sim N(0, 1)$. A la operación ilustrada en la ecuación 2-13 suele llamársele la **estandarización** de la variable aleatoria normal y. En la tabla I del apéndice se presenta la distribución normal estándar acumulada.

En muchas técnicas estadísticas se supone que la variable aleatoria sigue una distribución normal. El teorema del límite central es con frecuencia una justificación de la normalidad aproximada.

TEOREMA 2-1

El teorema del límite central

Si $y_1, y_2, \ldots, y_n$ es una sucesión de n variables aleatorias independientes que tienen una distribución idéntica con $E(y_i) = \mu$ y $V(y_i) = \sigma^2$ (ambas finitas) y $x = y_1 + y_2 + \ldots + y_n$, entonces

$$z_n = \frac{x - n\mu}{\sqrt{n\sigma^2}}$$

tiene una distribución $N(0, 1)$ aproximada en el sentido de que, si $F_n(z)$ es la función de la distribución de z_n y $\Phi(z)$ es la función de la distribución de la variable aleatoria $N(0, 1)$, entonces $\lim_{n \to \infty}[F_n(z)/\Phi(z)] = 1$.

Este resultado establece en esencia que la suma de n variables aleatorias independientes que tienen una distribución idéntica sigue una distribución aproximadamente normal. En muchos casos esta aproximación es adecuada para valores muy pequeños de n, digamos $n < 10$, mientras que en otros casos se necesita un valor grande de n, digamos $n > 100$. Frecuentemente se considera que el error de un experimento surge de una manera aditiva de varias fuentes independientes; por consiguiente, la distribución normal se convierte en un modelo recomendable para el error experimental combinado.

Una importante distribución de muestreo que puede definirse en términos de variables aleatorias normales es la **distribución χ^2 o ji-cuadrada**. Si $z_1, z_2, \ldots, z_k$ son variables aleatorias que tienen una distribución normal e independiente con media 0 y varianza 1, cuya abreviatura es NID(0, 1), entonces la variable aleatoria

$$x = z_1^2 + z_2^2 + \cdots + z_k^2$$

sigue la distribución ji-cuadrada con k grados de libertad. La función de densidad de la distribución ji-cuadrada es

$$f(x) = \frac{1}{2^{k/2}\Gamma\left(\frac{k}{2}\right)} x^{(k/2)-1} e^{-x/2} \qquad x > 0 \tag{2-14}$$

En la figura 2-6 se ilustran varias distribuciones ji-cuadrada. La distribución es asimétrica, o **sesgada**, con media y varianza

$$\mu = k$$
$$\sigma^2 = 2k$$

respectivamente. En la tabla III del apéndice se presentan los puntos porcentuales de la distribución ji-cuadrada.

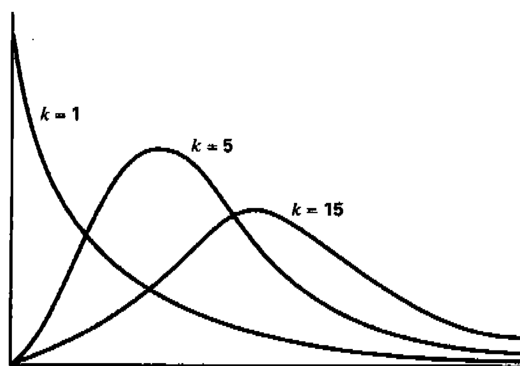


Figura 2-6 Varias distribuciones ji-cuadrada.

Como un ejemplo de una variable aleatoria que sigue la distribución ji-cuadrada, suponga que $y_1, y_2, \ldots, y_n$ es una muestra aleatoria de una distribución $N(\mu, \sigma^2)$. Entonces

$$\frac{SS}{\sigma^2} = \frac{\sum_{i=1}^{n} (y_i - \bar{y})^2}{\sigma^2} \sim \chi^2_{n-1} \tag{2-15}$$

Es decir, SS/σ^2 sigue una distribución ji-cuadrada con $n - 1$ grados de libertad.

Muchas de las técnicas utilizadas en este libro requieren el cálculo y la manipulación de sumas de cuadrados. El resultado dado en la ecuación 2-15 es de suma importancia y aparece en múltiples ocasiones; cuando una suma de cuadrados de variables aleatorias normales se divide por σ^2 sigue la distribución ji-cuadrada.

Al examinar la ecuación 2-8, se observa que la varianza muestral puede escribirse como

$$S^2 = \frac{SS}{n-1} \tag{2-16}$$

Si las observaciones de la muestra son $\text{NID}(\mu, \sigma^2)$, entonces la distribución de S^2 es $[\sigma^2/(n-1)]\chi^2_{n-1}$. Por lo tanto, la distribución de muestreo de la varianza muestral es una constante multiplicada por la distribución ji-cuadrada si la población tiene una distribución normal.

Si z y χ^2_k son variables aleatorias independientes normal estándar y ji-cuadrada, respectivamente, la variable aleatoria

$$t_k = \frac{z}{\sqrt{\chi^2_k / k}} \tag{2-17}$$

sigue la **distribución t con k grados de libertad**, denotada t_k. La función de densidad de t es

$$f(t) = \frac{\Gamma[(k+1)/2]}{\sqrt{k\pi}\,\Gamma(k/2)} \frac{1}{[(t^2/k)+1]^{(k+1)/2}} \qquad -\infty < t < \infty \tag{2-18}$$

y la media y la varianza de t son $\mu = 0$ y $\sigma^2 = k/(k-2)$ para $k > 2$, respectivamente. En la figura 2-7 se ilustran varias distribuciones t. Observe que si $k = \infty$, la distribución t se convierte en la distribución normal

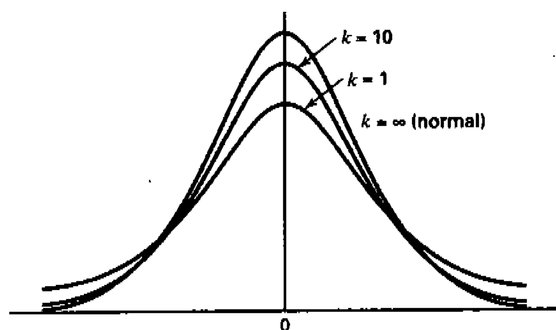


Figura 2-7 Varias distribuciones t.

estándar. En la tabla II del apéndice se presentan los puntos porcentuales de la distribución t. Si $y_1, y_2, \ldots, y_n$ es una muestra aleatoria de una distribución $N(\mu, \sigma^2)$, entonces la cantidad

$$t = \frac{\bar{y} - \mu}{S / \sqrt{n}} \tag{2-19}$$

se distribuye como t con $n - 1$ grados de libertad.

La última distribución de muestreo que consideraremos es la **distribución F**. Si χ_u^2 y χ_v^2 son dos variables aleatorias ji-cuadrada independientes con u y v grados de libertad, respectivamente, entonces el cociente

$$F_{u,v} = \frac{\chi_u^2 / u}{\chi_v^2 / v} \tag{2-20}$$

sigue la distribución F con u grados de libertad en el *numerador* y v grados de libertad en el *denominador*. Si x es una variable aleatoria F con u grados de libertad en el numerador y v grados de libertad en el denominador, entonces la distribución de probabilidad de x es

$$h(x) = \frac{\Gamma\left(\frac{u+v}{2}\right)\left(\frac{u}{v}\right)^{u/2} x^{(u/2)-1}}{\Gamma\left(\frac{u}{2}\right)\Gamma\left(\frac{v}{2}\right)\left[\left(\frac{u}{v}\right)x+1\right]^{(u+v)/2}} \qquad 0 < x < \infty \tag{2-21}$$

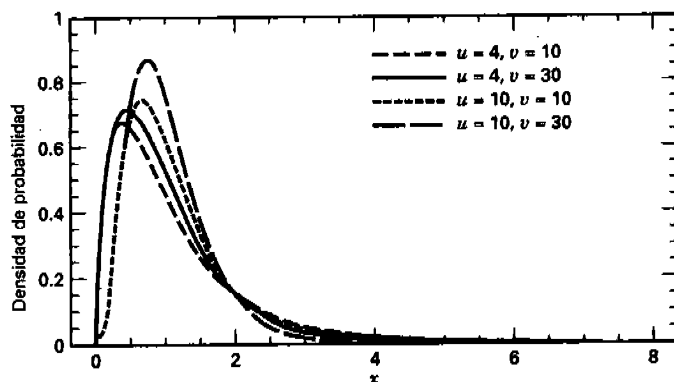


Figura 2-8 Varias distribuciones F.

En la figura 2-8 se ilustran varias distribuciones *F*. Esta distribución es muy importante en el análisis estadístico de experimentos diseñados. En la tabla IV del apéndice se presentan los puntos porcentuales de la distribución *F*.

Como un ejemplo de un estadístico que sigue una distribución *F*, suponga que se tienen dos poblaciones normales independientes con varianza común σ^2. Si $y_{11}, y_{12}, \ldots, y_{1n_1}$ es una muestra aleatoria de n_1 observaciones de la primera población y si $y_{21}, y_{22}, \ldots, y_{2n_2}$ es una muestra aleatoria de n_2 observaciones de la segunda, entonces

$$\frac{S_1^2}{S_2^2} \sim F_{n_1-1,\, n_2-1} \tag{2-22}$$

donde S_1^2 y S_2^2 son las dos varianzas muestrales. Este resultado se sigue directamente de las ecuaciones 2-15 y 2-20.

2-4 INFERENCIAS ACERCA DE LAS DIFERENCIAS EN LAS MEDIAS, DISEÑOS ALEATORIZADOS

Estamos preparados ahora para volver al problema del mortero de cemento portland de la sección 2-1. Recuerde que se estaban investigando dos formulaciones diferentes para determinar si difieren en la fuerza de la tensión de adhesión. En esta sección se examina cómo pueden analizarse los datos de este experimento comparativo simple utilizando procedimientos de **prueba de hipótesis** e **intervalos de confianza** para comparar las medias de dos tratamientos.

A lo largo de esta sección se supone que se usa un **diseño experimental completamente aleatorizado.** En este diseño, los datos se consideran como si fueran una muestra aleatoria de una distribución normal.

2-4.1 Prueba de hipótesis

Se retoma ahora el experimento del cemento portland introducido en la sección 2-1. Recuerde que el interés se encuentra en comparar la fuerza de dos formulaciones diferentes: una del mortero sin modificar y una del mortero modificado. En general, estas dos formulaciones pueden considerarse como dos **niveles del factor** "formulaciones". Sea que $y_{11}, y_{12}, \ldots, y_{1n_1}$ represente las n_1 observaciones del primer nivel del factor y que $y_{21}, y_{22}, \ldots, y_{2n_2}$ represente las n_2 observaciones del segundo nivel del factor. Se supone que las muestras se sacan al azar de dos poblaciones normales independientes. En la figura 2-9 se ilustra la situación.

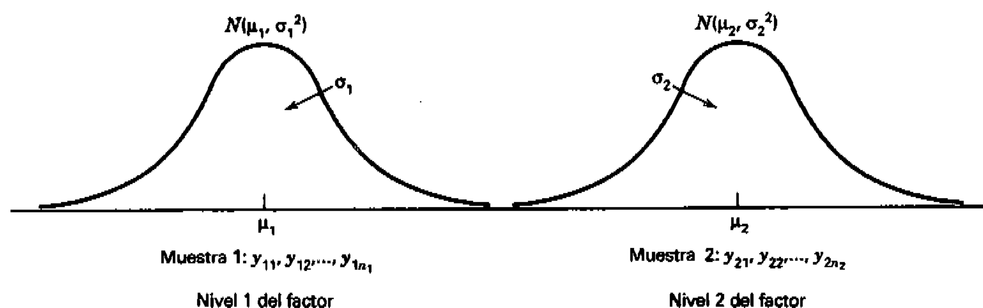


Figura 2-9 La situación del muestreo para la prueba *t* de dos muestras.

Un modelo de los datos

Con frecuencia los resultados de un experimento se describen con un **modelo**. Un modelo estadístico simple que describe los datos de un experimento como el que acaba de describirse es

$$y_{ij} = \mu_i + \varepsilon_{ij} \begin{cases} i = 1, 2 \\ j = 1, 2, \ldots, n_i \end{cases} \tag{2-23}$$

donde y_{ij} es la observación j-ésima del nivel i del factor, μ_i es la media de la respuesta para el nivel i-ésimo del factor, y ε_{ij} es una variable aleatoria normal asociada con la observación ij-ésima. Se supone que las ε_{ij} son NID$(0, \sigma_1^2)$, $i = 1, 2$. Se acostumbra hacer referencia a ε_{ij} como el componente del **error aleatorio** del modelo. Puesto que las medias μ_1 y μ_2 son constantes, se observa directamente a partir del modelo que las y_{ij} son NID(μ_i, σ_1^2), $i = 1, 2$, como se acaba de suponer arriba. Para más información acerca de los modelos de los datos, referirse al material suplementario del texto.

Hipótesis estadísticas

Una **hipótesis estadística** es un enunciado o afirmación ya sea acerca de los parámetros de una distribución de probabilidad o de los parámetros de un modelo. La hipótesis refleja alguna **conjetura** acerca de la situación del problema. Por ejemplo, en el experimento del cemento portland, puede pensarse que las fuerzas de la tensión de adhesión promedio de las dos formulaciones del mortero son iguales. Esto puede enunciarse formalmente como

$$H_0: \mu_1 = \mu_2$$
$$H_1: \mu_1 \neq \mu_2$$

donde μ_1 es la fuerza de la tensión de adhesión promedio del mortero modificado y μ_2 es la fuerza de tensión de enlace promedio del mortero sin modificar. Al enunciado $H_0{:}\mu_1 = \mu_2$ se le llama la **hipótesis nula** y a $H_1{:}\mu_1 \neq \mu_2$ se le llama la **hipótesis alternativa**. A la hipótesis alternativa que se especifica aquí se le llama **hipótesis alternativa de dos colas** porque sería verdadera si $\mu_1 < \mu_2$ o si $\mu_1 > \mu_2$.

Para probar una hipótesis se proyecta un procedimiento para tomar una muestra aleatoria, calcular un **estadístico de prueba** apropiado para después rechazar o no estar en posición de rechazar la hipótesis nula H_0. Parte de este procedimiento consiste en especificar el conjunto de valores del estadístico de prueba que llevan al rechazo de H_0. A este conjunto de valores se le llama la **región crítica** o **región de rechazo** de la prueba.

Pueden cometerse dos tipos de errores cuando se prueban hipótesis. Si la hipótesis nula se rechaza cuando es verdadera, ha ocurrido un error tipo I. Si la hipótesis nula *no* se rechaza cuando es falsa, se ha cometido un error tipo II. Las probabilidades de estos dos errores se expresan con símbolos especiales:

$$\alpha = P(\text{error tipo I}) = P(\text{rechazar } H_0 \mid H_0 \text{ es verdadera})$$
$$\beta = P(\text{error tipo II}) = P(\text{dejar de rechazar } H_0 \mid H_0 \text{ es falsa})$$

En ocasiones es más conveniente trabajar con la **potencia** de la prueba, donde

$$\text{Potencia} = 1 - \beta = P(\text{rechazar } H_0 \mid H_0 \text{ es falsa})$$

El procedimiento general en la prueba de hipótesis es especificar un valor de la probabilidad α del error tipo I, llamada con frecuencia el **nivel de significación** de la prueba, y después diseñar el procedimiento de prueba de tal modo que la probabilidad β del error tipo II tenga un valor convenientemente pequeño.

La prueba t de dos muestras

Considere que puede suponerse que las varianzas de las fuerzas de la tensión de adhesión fueron idénticas para ambas formulaciones del mortero. Entonces el estadístico de prueba que deberá usarse para comparar las medias de dos tratamientos en el diseño completamente aleatorizado es

$$t_0 = \frac{\bar{y}_1 - \bar{y}_2}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \tag{2-24}$$

donde $\bar{y}_1$ y $\bar{y}_2$ son las medias muestrales, n_1 y n_2 son los tamaños de las muestras, S_p^2 es una estimación de la varianza común $\sigma_1^2 = \sigma_2^2 = \sigma^2$ calculada a partir de

$$S_p^2 = \frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2} \tag{2-25}$$

y S_1^2 y S_2^2 son las dos varianzas muestrales individuales. Para determinar si deberá rechazarse $H_0{:}\mu_1 = \mu_2$, se compararía t_0 con la distribución t con $n_1 + n_2 - 2$ grados de libertad. Si $|t_0| > t_{\alpha/2,n_1+n_2-2}$, donde $t_{\alpha/2,n_1+n_2-2}$ es el punto porcentual $\alpha/2$ superior de la distribución t con $n_1 + n_2 - 2$ grados de libertad, entonces se *rechazaría* H_0 y se concluiría que las fuerzas promedio de las dos formulaciones del mortero de cemento portland difieren. A este procedimiento de prueba se le llama generalmente la **prueba t de dos muestras**.

Este procedimiento puede justificarse de la siguiente manera. Si el muestreo se está haciendo de distribuciones normales independientes, entonces la distribución de $\bar{y}_1 - \bar{y}_2$ es $N[\mu_1 - \mu_2, \sigma^2(1/n_1 + 1/n_2)]$. Por lo tanto, si se conociera σ^2, y si $H_0{:}\mu_1 = \mu_2$ fuera verdadera, la distribución de

$$Z_0 = \frac{\bar{y}_1 - \bar{y}_2}{\sigma\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \tag{2-26}$$

sería $N(0, 1)$. Sin embargo, al sustituir σ con S_p en la ecuación 2-26, la distribución de Z_0 cambia de la normal estándar a la distribución t con $n_1 + n_2 - 2$ grados de libertad. Ahora bien, si H_0 es verdadera, t_0 de la ecuación 2-24 se distribuye como $t_{n_1+n_2-2}$ y, por consiguiente, se esperaría que $100(1-\alpha)$ por ciento de los valores de t_0 estén entre $-t_{\alpha/2,n_1+n_2-2}$ y $t_{\alpha/2,n_1+n_2-2}$. Una muestra que produjera un valor de t_0 que estuviera fuera de estos límites sería inusual si la hipótesis nula fuera verdadera y es evidencia de que H_0 deberá rechazarse. Por lo tanto, la distribución t con $n_1 + n_2 - 2$ grados de libertad es la **distribución de referencia** apropiada para el estadístico de prueba t_0. Es decir, describe el comportamiento de t_0 cuando la hipótesis nula es verdadera. Observe que α es la probabilidad del error tipo I de la prueba.

En algunos problemas quizá quiera rechazarse H_0 únicamente si una de las medias es mayor que la otra. Por lo tanto, se especificaría una **hipótesis alternativa de una cola** $H_1{:}\mu_1 > \mu_2$ y H_0 sólo se rechazaría si $t_0 > t_{\alpha,n_1+n_2-2}$. Si se desea rechazar H_0 sólo si μ_1 es menor que μ_2, entonces la hipótesis alternativa es $H_1{:}\mu_1 < \mu_2$, y H_0 se rechazaría si $t_0 < -t_{\alpha,n_1+n_2-2}$.

Para ilustrar el procedimiento, considere los datos del cemento portland de la tabla 2-1. Para estos datos, se encuentra que

Mortero modificado	**Mortero sin modificar**
$\bar{y}_1 = 16.76$ kgf / cm^2	$\bar{y}_2 = 17.92$ kgf / cm^2
$S_1^2 = 0.100$	$S_2^2 = 0.061$
$S_1 = 0.316$	$S_2 = 0.247$
$n_1 = 10$	$n_2 = 10$

Puesto que las desviaciones estándar muestrales son razonablemente similares, no es improcedente concluir que las desviaciones estándar (o las varianzas) poblacionales son iguales. Por lo tanto, puede usarse la ecuación 2-24 para probar las hipótesis

$$H_0: \mu_1 = \mu_2$$
$$H_1: \mu_1 \neq \mu_2$$

Además, $n_1 + n_2 - 2 = 10 + 10 - 2 = 18$, y si se elige $\alpha = 0.05$, entonces $H_0{:}\mu_1 = \mu_2$ se rechazaría si el valor numérico del estadístico de prueba $t_0 > t_{0.025,\,18} = 2.101$, o si $t_0 < -t_{0.025,\,18} = -2.101$. Estos límites de la región crítica se ilustran en la distribución de referencia (t con 18 grados de libertad) de la figura 2-10.

Al utilizar la ecuación 2-25 se encuentra que

$$\begin{aligned} S_p^2 &= \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} \\ &= \frac{9(0.100) + 9(0.061)}{10 + 10 - 2} \\ &= 0.081 \\ S_p &= 0.284 \end{aligned}$$

y el estadístico de prueba es

$$\begin{aligned} t_0 &= \frac{\bar{y}_1 - \bar{y}_2}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \\ &= \frac{16.76 - 17.92}{0.284\sqrt{\frac{1}{10} + \frac{1}{10}}} \\ &= -9.13 \end{aligned}$$

Puesto que $t_0 = -9.13 < -t_{0.025,\,18} = -2.101$, se rechazaría H_0 y se concluiría que las fuerzas de la tensión de adhesión promedio de las dos formulaciones del mortero de cemento portland son diferentes.

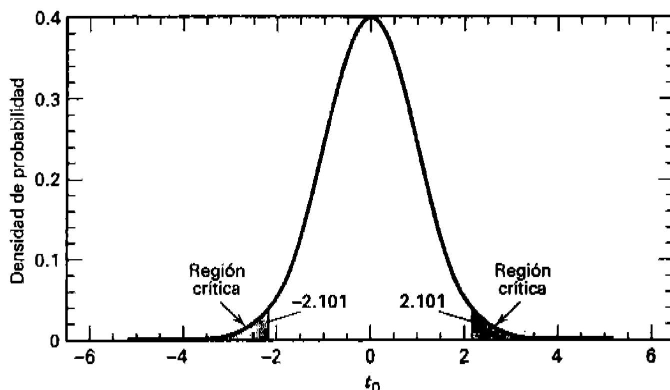


Figura 2-10 La distribución t con 18 grados de libertad con la región crítica $\pm t_{0.025,\,18} = \pm 2.101$.

El uso de valores P en la prueba de hipótesis

Una manera de reportar los resultados de una prueba de hipótesis es estableciendo que la hipótesis nula fue rechazada o no para un valor de α o **nivel de significación** específico. Por ejemplo, en el experimento del mortero de cemento portland anterior puede decirse que $H_0{:}\mu_1 = \mu_2$ se rechazó con el nivel de significación 0.05. Esta enunciación de las conclusiones es con frecuencia inadecuada porque no le ofrece al responsable de la toma de decisiones idea alguna de si el valor calculado del estadístico de prueba apenas rebasó la región de rechazo o si se adentró bastante en la misma. Además, al darse los resultados de esta manera se les impone a otros usuarios de la información el nivel de significación predefinido. Este enfoque puede ser insatisfactorio porque algunos responsables de la toma de decisiones podrían sentirse incómodos con los riesgos que implica el valor $\alpha = 0.05$.

Para evitar estas dificultades, en la práctica se ha adoptado extensivamente el **enfoque del valor *P*.** El valor *P* es la probabilidad de que el estadístico de prueba asuma un valor que sea al menos tan extremo como el valor observado del estadístico cuando la hipótesis nula H_0 es verdadera. Por lo tanto, un valor *P* transmite mucha información acerca del peso de la evidencia en contra de H_0 y, por consiguiente, el responsable de la toma de decisiones puede llegar a una conclusión con cualquier nivel de significación especificado. En términos más formales, el valor *P* se define como el nivel de significación menor que llevaría a rechazar la hipótesis nula H_0.

Se acostumbra decir que el estadístico de prueba (y los datos) es significativo cuando se rechaza la hipótesis nula; por lo tanto, el valor *P* puede considerarse como el menor nivel α en el que los datos son significativos. Una vez que se conoce el valor *P*, el responsable de la toma de decisiones puede determinar la medida en que los datos son significativos sin que el analista de los datos imponga formalmente un nivel de significación preseleccionado.

No siempre es sencillo calcular el valor *P* exacto de una prueba. Sin embargo la mayoría de los programas de computación modernos para realizar análisis estadísticos reportan valores *P*, y pueden obtenerse también en algunas calculadoras portátiles. A continuación se indicará cómo obtener una aproximación del valor *P* para el experimento del mortero de cemento portland. Por la tabla II del apéndice, para una distribución *t* con 18 grados de libertad, la probabilidad menor en el área de la cola es 0.0005, para la cual $t_{0.0005,18} = 3.922$. Ahora bien, $|t_0| = 9.13 > 3.922$, de donde, ya que la hipótesis alternativa es de dos colas, se sabe que el valor *P* debe ser menor que $2(0.0005) = 0.001$. Algunas calculadoras portátiles tienen la capacidad para calcular valores *P*. Una de ellas es la HP-48. Utilizando esta calculadora se obtiene el valor *P* para el valor $t_0 = -9.13$ del experimento de la formulación del mortero de cemento portland como $P = 3.68 \times 10^{-8}$. Por lo tanto, la hipótesis nula $H_0{:}\mu_1 = \mu_2$ se rechazaría con cualquier nivel de significación $\alpha \geq 3.68 \times 10^{-8}$.

Solución por computadora

Hay muchos paquetes de software de estadística que cuentan con la capacidad para probar hipótesis estadísticas. En la tabla 2-2 se presenta la salida del procedimiento para la prueba *t* de dos muestras de Minitab aplicado al experimento de la formulación del mortero de cemento portland. Observe que la salida incluye algunos estadísticos concisos acerca de las dos muestras (la abreviatura "SE Mean" ["SE media"] se refiere al error estándar de la media, $s / \sqrt{n}$), así como alguna información sobre los intervalos de confianza para la diferencia en las dos medias (los cuales se revisan en las secciones 2-4.3 y 2-6). El programa también prueba la hipótesis de interés, permitiendo que el analista especifique la naturaleza de la hipótesis alternativa ("not =" ["no ="] significa $H_1{:}\mu_1 \neq \mu_2$) y la elección de α ($\alpha = 0.05$ en este caso).

La salida incluye el valor calculado de t_0, el valor *P* (llamado el nivel de significación) y la decisión que debería tomarse dado el valor especificado de α. Observe que el valor calculado del estadístico *t* difiere ligeramente del valor que se calculó manualmente aquí y que el valor *P* que se reporta es $P = 0.0000$. Mu-

Tabla 2-2 Prueba *t* de dos muestras usando Minitab para el experimento del mortero de cemento portland

Prueba *t* de dos muestras e intervalo de confianza

```
Two sample T for Modified vs Unmod

              N        Mean        StDev       SE Mean
Modified     10      16.774        0.309         0.098
Unmod        10      17.922        0.248         0.078

95% CI for mu Modified - mu Unmod: ( -1.411,  -0.885)
t-Test mu Modified = mu Unmod (vs not =): T = -9.16
P = 0.0000 DF = 18
Both use Pooled StDev = 0.280
```

chos paquetes de software no reportarán un valor P real menor que 0.0001 y en su lugar presentan un valor "por omisión". Éste es el caso aquí.

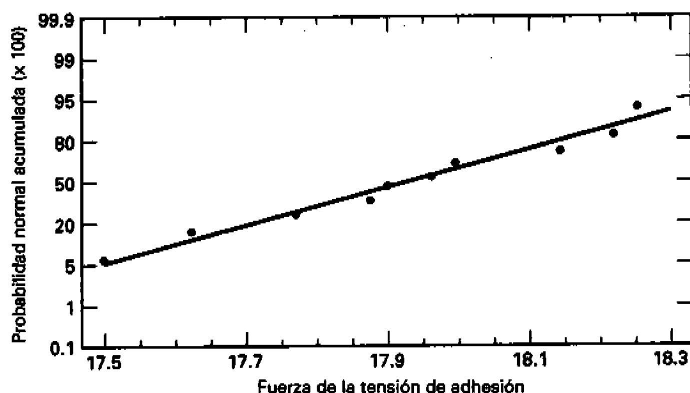
Verificación de los supuestos en la prueba *t*

Para utilizar el procedimiento de la prueba t se establecen los supuestos de que ambas muestras se toman de poblaciones independientes que pueden describirse con una distribución normal, que las desviaciones estándar o las varianzas de ambas poblaciones son iguales, y que las observaciones son variables aleatorias independientes. El supuesto de independencia es crítico, pero si el orden de las corridas está aleatorizado (y, de ser apropiado, se seleccionan al azar otras unidades y materiales experimentales), este supuesto por lo general se satisfará. Los supuestos de la igualdad de las varianzas y la normalidad son fáciles de verificar utilizando una **gráfica de probabilidad normal.**

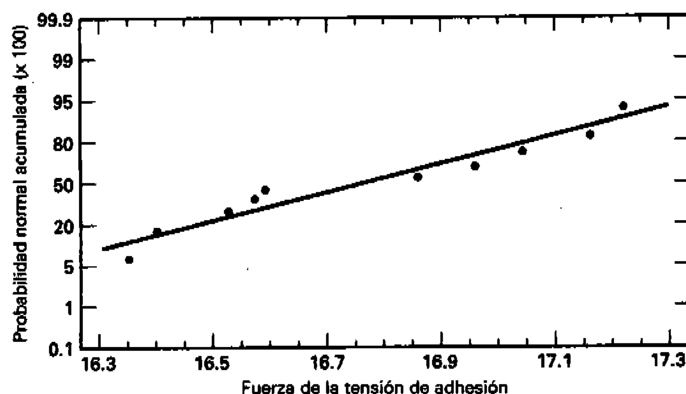
En general, la graficación de probabilidades es una técnica para determinar si los datos muestrales se ajustan a una distribución hipotetizada con base en un examen visual subjetivo de los datos. El procedimiento general es muy simple y puede realizarse rápidamente con la mayoría de los paquetes de software de estadística. En el material suplementario del texto se analiza la construcción manual de las gráficas de probabilidad normal.

Para construir una gráfica de probabilidad, primero se ordenan de menor a mayor las observaciones de la muestra. Es decir, la muestra $y_1, y_2, \ldots, y_n$ se ordena como $y_{(1)}, y_{(2)}, \ldots, y_{(n)}$, donde $y_{(1)}$ es la observación menor, $y_{(2)}$ es la segunda observación menor, y así sucesivamente, con $y_{(n)}$ la mayor. Las observaciones ordenadas $y_{(j)}$ se grafican entonces contra sus respectivas frecuencias acumuladas observadas $(j - 0.5)/n$. La escala de la frecuencia acumulada se ha dispuesto de tal modo que si la distribución hipotetizada describe de manera adecuada los datos, los puntos graficados estarán aproximadamente sobre una línea recta; si los puntos graficados muestran una desviación significativa de una recta, el modelo hipotetizado no es apropiado. Generalmente, determinar si los datos graficados pertenecen o no a una recta es una decisión subjetiva.

Para ilustrar el procedimiento, suponga que quiere verificarse el supuesto de que la fuerza de la tensión de adhesión en el experimento de la formulación del mortero de cemento portland sigue una distribución normal. Inicialmente sólo se consideran las observaciones de la formulación del mortero sin modificar. En la figura 2-11*a* se ilustra una gráfica de probabilidad normal generada por computadora. La mayoría de las gráficas de probabilidad normal muestran $100(j - 0.5)/n$ en la escala vertical izquierda (y en ocasiones se muestra $100[1 - (j - 0.5)/n]$ en la escala vertical derecha), con el valor de la variable graficado en la escala horizontal. Algunas gráficas de probabilidad normal convierten la frecuencia acumulada en un valor z normalizado. Una línea recta, elegida de manera subjetiva, se ha trazado en medio de los pun-



a) Mortero sin modificar



b) Mortero modificado

Figura 2-11 Gráficas de probabilidad normal de la fuerza de la tensión de adhesión en el experimento del cemento portland.

tos graficados. Al trazar la línea recta, uno deberá guiarse más por los puntos de la parte media de la gráfica que por los puntos extremos. Una buena regla empírica es trazar la recta aproximadamente entre los puntos de los cuartiles 25 y 75. Así se determinó la recta de la figura 2-11*a*. Para evaluar la "proximidad" de los puntos a la línea recta, imagine un lápiz grueso colocado sobre la recta. Si este lápiz imaginario cubre todos los puntos, entonces una distribución normal describe de manera adecuada los datos. Puesto que los puntos de la figura 2-11*a* pasarían la prueba del lápiz grueso, se concluye que la distribución normal es un modelo apropiado para la fuerza de la tensión de adhesión del mortero sin modificar. En la figura 2-11*b* se presenta la gráfica de probabilidad normal para las 10 observaciones de la fuerza de la tensión de adhesión del mortero modificado. De nueva cuenta, se concluiría que es razonable el supuesto de una distribución normal.

Es posible obtener una estimación de la media y la desviación estándar directamente de la gráfica de probabilidad normal. La media se estima como el percentil 50 de la gráfica de probabilidad y la desviación estándar se estima como la diferencia entre los percentiles 84 y 50. Esto significa que el supuesto de la igualdad de las varianzas poblacionales en el experimento del cemento portland puede verificarse comparando las pendientes de las dos rectas de las figuras 2-11*a* y 2-11*b*. Ambas rectas tienen pendientes muy si-

milares, por lo que el supuesto de la igualdad de las varianzas es razonable. Si se viola este supuesto, deberá usarse la versión de la prueba t que se describe en la sección 2-4.4. En el material suplementario del texto hay más información acerca de la verificación de los supuestos de la prueba t.

Cuando ocurren violaciones importantes de los supuestos, se afectará el desempeño de la prueba t. En general, las violaciones de pequeñas a moderadas no son motivo de preocupación particular, pero no deberá ignorarse *cualquier* falla del supuesto de independencia, así como los indicios claros de que no se satisface el supuesto de normalidad. Tanto el nivel de significación de la prueba como la capacidad para detectar diferencias entre las medias serán afectados adversamente por el incumplimiento de estos supuestos. Un recurso para resolver este problema son las **transformaciones**. Este tema se analiza con mayor detalle en el capítulo 3. También es posible utilizar procedimientos no paramétricos para la prueba de hipótesis cuando las observaciones provienen de poblaciones no normales. Referirse a Montgomery y Runger [83d] para más detalles.

Una justificación alternativa de la prueba *t*

La prueba t de dos muestras que acaba de presentarse depende en teoría del supuesto fundamental de que las dos poblaciones de las que se seleccionaron las muestras al azar son normales. Aun cuando el supuesto de normalidad es necesario para desarrollar formalmente el procedimiento de prueba, como ya se mencionó, las desviaciones moderadas de la normalidad no afectarán seriamente los resultados. Puede argumentarse (por ejemplo, ver Box, Hunter y Hunter [18]) que el uso de un diseño aleatorizado permite probar hipótesis sin *ningún* supuesto respecto de la forma de la distribución. En resumen, el razonamiento es el siguiente. Si los tratamientos no tienen ningún efecto, todas las [20!/(10!10!)] = 184,756 formas posibles en que podrían ocurrir las 20 observaciones son igualmente posibles. Hay un valor de t_0 para cada uno de estos 184,756 posibles arreglos. Si el valor de t_0 que se obtiene en realidad de los datos es inusualmente grande o inusualmente pequeño con referencia al conjunto de los 184,756 valores posibles, es una indicación de que $\mu_1 \neq \mu_2$.

A este tipo de procedimiento se le llama **prueba de aleatorización.** Puede demostrarse que la prueba t es una buena aproximación de la prueba de aleatorización. Por lo tanto, se usarán aquí pruebas t (y otros procedimientos que pueden considerarse aproximaciones de pruebas de aleatorización) sin prestar demasiada atención al supuesto de normalidad. Ésta es una de las razones por las que un procedimiento simple, como las gráficas de probabilidad normal, es adecuado para verificar el supuesto de normalidad.

2-4.2 Elección del tamaño de la muestra

La elección de un tamaño de la muestra apropiado es uno de los aspectos más importantes de cualquier problema de diseño experimental. La elección del tamaño de la muestra y la probabilidad β del error tipo II guardan una estrecha relación. Suponga que se están probando las hipótesis

$$H_0: \mu_1 = \mu_2$$
$$H_1: \mu_1 \neq \mu_2$$

y que las medias no son iguales, por lo que $\delta = \mu_1 - \mu_2$. Puesto que $H_0: \mu_1 = \mu_2$ no es verdadera, la preocupación principal es cometer la equivocación de no rechazar H_0. La probabilidad del error tipo II depende de la verdadera diferencia en las medias δ. A una gráfica de β contra δ para un tamaño particular de la muestra se le llama la **curva de operación característica,** o **curva OC**, de la prueba. El error β también es una función del tamaño de la muestra. En general, para un valor dado de δ, el error β se reduce cuando el tamaño de la muestra se incrementa. Es decir, es más fácil detectar una diferencia especificada en las medias para tamaños grandes de la muestra que para los tamaños pequeños.

En la figura 2-12 se muestra un juego de curvas de operación característica para las hipótesis

$$H_0: \mu_1 = \mu_2$$
$$H_1: \mu_1 \neq \mu_2$$

para el caso en que las dos varianzas poblacionales σ_1^2 y σ_2^2 son desconocidas pero iguales ($\sigma_1^2 = \sigma_2^2 = \sigma^2$) y para un nivel de significación de $\alpha = 0.05$. Las curvas también parten del supuesto de que los tamaños de las muestras de las dos poblaciones son iguales; es decir, $n_1 = n_2 = n$. El parámetro del eje horizontal de la figura 2-12 es

$$d = \frac{|\mu_1 - \mu_2|}{2\sigma} = \frac{|\delta|}{2\sigma}$$

La división de $|\delta|$ por 2σ permite al experimentador usar el mismo juego de curvas, independientemente del valor de la varianza (la diferencia en las medias se expresa en unidades de desviación estándar). Por otra parte, el tamaño de la muestra usado para construir las curvas es en realidad $n^* = 2n - 1$.

Al examinar estas curvas, se observa lo siguiente:

1. Entre más grande sea la diferencia en las medias, $\mu_1 - \mu_2$, menor será la probabilidad del error tipo II para un tamaño de la muestra y un valor de α dados. Es decir, para un tamaño de la muestra y un valor de α especificados, la prueba detectará con mayor facilidad las diferencias grandes que las pequeñas.
2. Cuando el tamaño de la muestra se hace más grande, la probabilidad del error tipo II se hace más pequeña para una diferencia en las medias y un valor de α dados. Es decir, para detectar una diferencia δ especificada, puede aumentarse la potencia de la prueba incrementando el tamaño de la muestra.

Las curvas de operación característica son con frecuencia útiles para seleccionar el tamaño de la muestra que debe usarse en un experimento. Por ejemplo, considere el problema del mortero de cemento portland comentado antes. Suponga que si las dos formulaciones difieren en la fuerza promedio hasta en 0.5 kgf/cm^2, sería deseable detectarlo con una probabilidad alta. Por lo tanto, puesto que $\mu_1 - \mu_2 =$

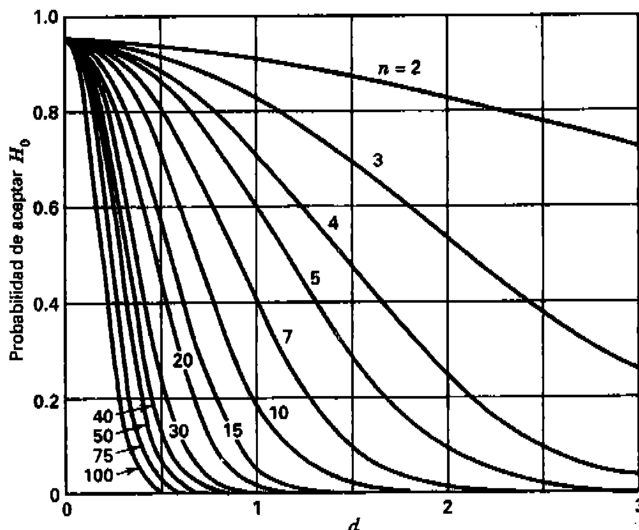


Figura 2-12 Curvas de operación característica para la prueba t de dos colas con $\alpha = 0.05$. (Reproducida con permiso de "Operating Characteristics Curves for the Common Statistical Tests of Significance", C.L. Ferris, F.E. Grubbs y C.L. Weaver, *Annals of Mathematical Statistics*.)

0.5 kgf/cm^2 es la diferencia "crítica" en las medias que quiere detectarse, se encuentra que d, el parámetro del eje horizontal de la curva de operación característica de la figura 2-12, es

$$d = \frac{|\mu_1 - \mu_2|}{2\sigma} = \frac{0.5}{2\sigma} = \frac{0.25}{\sigma}$$

Desafortunadamente, d incluye al parámetro desconocido σ. Sin embargo, suponga que con base en la experiencia previa se piensa que es altamente improbable que la desviación estándar de cualquiera de las observaciones de la fuerza exceda 0.25 kgf/cm^2. Entonces al usar $\sigma = 0.25$ en la expresión anterior para d se obtiene $d = 1$. Si quiere rechazarse la hipótesis nula 95% de las veces cuando $\mu_1 - \mu_2 = 0.5$, entonces $\beta = 0.05$, y en la figura 2-12 con $\beta = 0.05$ y $d = 1$ se obtiene $n^* = 16$, aproximadamente. Por lo tanto, puesto que $n^* = 2n - 1$, el tamaño de la muestra requerido es

$$n = \frac{n^* + 1}{2} = \frac{16+1}{2} = 8.5 \simeq 9$$

y se usarían los tamaños de las muestras $n_1 = n_2 = n = 9$.

En el ejemplo que se ha venido considerando, el experimentador utilizó en realidad un tamaño de la muestra de 10. Quizás el experimentador decidió incrementar ligeramente el tamaño de la muestra a fin de prevenir la posibilidad de que la estimación previa de la desviación estándar común σ haya sido demasiado conservadora y quizá fuera un poco mayor que 0.25.

Las curvas de operación característica desempeñan con frecuencia un papel importante en la elección del tamaño de la muestra en los problemas de diseño experimental. Su utilización a este respecto se revisa en capítulos subsecuentes. Para un análisis de los usos de las curvas de operación característica en otros experimentos comparativos simples similares a la prueba t de dos muestras, véase Montgomery y Runger [83d].

2-4.3 Intervalos de confianza

Aun cuando la prueba de hipótesis es un procedimiento útil, en ocasiones no cuenta la historia completa. Muchas veces es preferible proporcionar un intervalo dentro del cual cabría esperar que estaría incluido el valor del parámetro o los parámetros en cuestión. A las declaraciones de estos intervalos se les llama **intervalos de confianza**. En muchos experimentos de ingeniería e industriales, el experimentador sabe de antemano que las medias μ_1 y μ_2 difieren; por consiguiente, la prueba de la hipótesis $\mu_1 = \mu_2$ es de escaso interés. Por lo general el experimentador estaría más interesado en un intervalo de confianza para la diferencia en las medias $\mu_1 - \mu_2$.

Para definir un intervalo de confianza, suponga que θ es un parámetro desconocido. Para obtener una estimación del intervalo de θ, es necesario encontrar dos estadísticos L y U tales que la declaración de probabilidad

$$P(L \leq \theta \leq U) = 1 - \alpha \qquad (2\text{-}27)$$

sea verdadera. Al intervalo

$$L \leq \theta \leq U \qquad (2\text{-}28)$$

se le llama **intervalo de confianza de 100(1 – α) por ciento** para el parámetro θ. La interpretación de este intervalo es que si, en muestreos aleatorios repetidos, se construye gran número de estos intervalos, 100(1 – α) por ciento de ellos contendrán el verdadero valor de θ. A los estadísticos L y U se les llama los **límites**

de confianza inferior y **superior**, respectivamente, y a $1 - \alpha$ se le llama el **coeficiente de confianza**. Si $\alpha = 0.05$, a la ecuación 2-28 se le llama intervalo de confianza de 95% para θ. Observe que los intervalos de confianza tienen una interpretación de frecuencia; es decir, no se sabe si la declaración es verdadera para esta muestra específica, pero sí se sabe que el método usado para generar el intervalo de confianza produce declaraciones correctas en $100(1 - \alpha)$ por ciento de las veces.

Suponga que quiere encontrarse un intervalo de confianza de $100(1-\alpha)$ por ciento para la verdadera diferencia de las medias $\mu_1 - \mu_2$ en el problema del cemento portland. El intervalo puede deducirse de la siguiente manera. El estadístico

$$\frac{\bar{y}_1 - \bar{y}_2 - (\mu_1 - \mu_2)}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

se distribuye como $t_{n_1+n_2-2}$. Por lo tanto,

$$P\left(-t_{\alpha/2,n_1+n_2-2} \le \frac{\bar{y}_1 - \bar{y}_2 - (\mu_1 - \mu_2)}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \le t_{\alpha/2,n_1+n_2-2}\right) = 1 - \alpha$$

o

$$P\left(\bar{y}_1 - \bar{y}_2 - t_{\alpha/2,n_1+n_2-2} S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \le \mu_1 - \mu_2 \le \bar{y}_1 - \bar{y}_2 + t_{\alpha/2,n_1+n_2-2} S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\right) = 1 - \alpha \qquad (2\text{-}29)$$

Al comparar las ecuaciones 2-29 y 2-27, se observa que

$$\bar{y} - \bar{y}_2 - t_{\alpha/2,n_1+n_2-2} S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \le \mu_1 - \mu_2 \le \bar{y}_1 - \bar{y}_2 + t_{\alpha/2,n_1+n_2-2} S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \qquad (2\text{-}30)$$

es un intervalo de confianza de $100(1 - \alpha)$ por ciento para $\mu_1 - \mu_2$.

La estimación real del intervalo de confianza de 95% para la diferencia en la fuerza de la tensión de adhesión promedio de las formulaciones del mortero de cemento portland se encuentra haciendo la siguiente sustitución en la ecuación 2-30:

$$16.76 - 17.92 - (2.101)0.284\sqrt{\tfrac{1}{10} + \tfrac{1}{10}} \le \mu_1 - \mu_2 \le 16.76 - 17.92 + (2.101)0.284\sqrt{\tfrac{1}{10} + \tfrac{1}{10}}$$

$$-1.16 - 0.27 \le \mu_1 - \mu_2 \le -1.16 + 0.27$$

$$-1.43 \le \mu_1 - \mu_2 \le -0.89$$

Por lo tanto, el intervalo de confianza de 95% estimado para la diferencia en las medias se extiende de -1.43 kgf/cm^2 a -0.89 kgf/cm^2. Expresado en otros términos, el intervalo de confianza es $\mu_1 - \mu_2 = -1.16$ kgf/cm^2 $\pm$ 0.27 kgf/cm^2, o la diferencia en las fuerzas promedio es -1.16 kgf/cm^2, y la precisión de esta estimación es de $\pm$0.27 kgf/cm^2. Observe que como $\mu_1 - \mu_2 = 0$ *no* está incluida en este intervalo, los datos no apoyan la hipótesis de que $\mu_1 = \mu_2$ con el nivel de significación de 5%. Es probable que la fuerza media de la formulación sin modificar exceda la fuerza media de la formulación modificada. Observe que en la ta-

bla 2-2 Minitab también reportó este intervalo de confianza cuando se llevó a cabo el procedimiento de la prueba de hipótesis.

2-4.4 Caso en que $\sigma_1^2 \neq \sigma_2^2$

Si se está probando

$$H_0: \mu_1 = \mu_2$$
$$H_1: \mu_1 \neq \mu_2$$

y no hay bases para suponer que las varianzas σ_1^2 y σ_2^2 son iguales, entonces es necesario hacer ligeras modificaciones en la prueba t de dos muestras. En este caso el estadístico de prueba es

$$t_0 = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \tag{2-31}$$

Este estadístico no se distribuye exactamente como t. No obstante, t es una buena aproximación de la distribución de t_0 si se usa

$$v = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{(S_1^2/n_1)^2}{n_1 - 1} + \frac{(S_2^2/n_2)^2}{n_2 - 1}} \tag{2-32}$$

para los grados de libertad. Una indicación clara de la desigualdad de las varianzas en una gráfica de probabilidad normal sería una situación que requeriría esta versión de la prueba t. El lector no deberá encontrar problemas para desarrollar una ecuación para encontrar ese intervalo de confianza para la diferencia en las medias en el caso de varianzas desiguales.

2-4.5 Caso en que se conocen σ_1^2 y σ_2^2

Si las varianzas de ambas poblaciones *se conocen*, entonces las hipótesis

$$H_0: \mu_1 = \mu_2$$
$$H_1: \mu_1 \neq \mu_2$$

pueden probarse utilizando el estadístico

$$Z_0 = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \tag{2-33}$$

Si ambas poblaciones son normales, o si los tamaños de las muestras son lo suficientemente grandes para aplicar el teorema del límite central, la distribución de Z_0 es $N(0, 1)$ si la hipótesis nula es verdadera. Por lo tanto, la región crítica se encontraría utilizando la distribución normal en lugar de la distribución t. Específicamente, H_0 se rechazaría si $|Z_0| > Z_{\alpha/2}$, donde $Z_{\alpha/2}$ es el punto porcentual $\alpha/2$ superior de la distribución normal estándar.

A diferencia de la prueba t de las secciones anteriores, en la prueba de las medias con varianzas conocidas no se requiere el supuesto de que el muestreo se haga de poblaciones normales. Puede aplicarse el teorema del límite central para justificar una distribución normal aproximada para la diferencia en las medias muestrales $\bar{y}_1 - \bar{y}_2$.

El intervalo de confianza de $100(1-\alpha)$ por ciento para $\mu_1 - \mu_2$ cuando las varianzas se conocen es

$$\bar{y}_1 - \bar{y}_2 - Z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}} \leq \mu_1 - \mu_2 \leq \bar{y}_1 - \bar{y}_2 + Z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}} \tag{2-34}$$

Como ya se señaló, el intervalo de confianza es con frecuencia un complemento útil del procedimiento de prueba de hipótesis.

2-4.6 Comparación de una sola media con un valor especificado

Algunos experimentos incluyen la comparación de la media μ de una sola población con un valor especificado, por ejemplo μ_0. Las hipótesis son

$$H_0: \mu = \mu_0$$
$$H_1: \mu \neq \mu_0$$

Si la población es normal con varianza conocida, o si la población no es normal pero el tamaño de la muestra es lo suficientemente grande para aplicar el teorema del límite central, entonces la hipótesis puede probarse utilizando una aplicación directa de la distribución normal. El estadístico de prueba es

$$Z_0 = \frac{\bar{y} - \mu_0}{\sigma / \sqrt{n}} \tag{2-35}$$

Si $H_0{:}\mu = \mu_0$ es verdadera, entonces la distribución de Z_0 es $N(0, 1)$. Por lo tanto, la regla de decisión para $H_0{:}\mu = \mu_0$ es rechazar la hipótesis nula si $|Z_0| > Z_{\alpha/2}$. El valor de la media μ_0 especificado en la hipótesis nula se determina por lo general mediante una de las tres formas siguientes. Puede ser resultado de evidencia, conocimientos o experimentación previos. Puede ser resultado de alguna teoría o modelo que describe la situación bajo estudio. Por último, puede ser resultado de especificaciones contractuales.

El intervalo de confianza de $100(1 - \alpha)$ por ciento para la verdadera media poblacional es

$$\bar{y} - Z_{\alpha/2}\sigma / \sqrt{n} \leq \mu \leq \bar{y} + Z_{\alpha/2}\sigma / \sqrt{n} \tag{2-36}$$

EJEMPLO 2-1 ....

Un proveedor ofrece lotes de tela a un fabricante de textiles. El fabricante desea saber si la resistencia a la ruptura promedio excede 200 psi. De ser así, el fabricante aceptará el lote. La experiencia pasada indica que un valor razonable para la varianza de la resistencia a la ruptura es $100(\text{psi})^2$. Las hipótesis que deberán probarse son

$$H_0: \mu = 200$$
$$H_1: \mu > 200$$

Observe que se trata de una hipótesis alternativa de una cola. Por lo tanto, el lote se aceptaría sólo si la hipótesis nula $H_0{:}\mu = 200$ pudiera rechazarse (es decir, si $Z_0 > Z_\alpha$).

Se seleccionan cuatro ejemplares aleatoriamente, y la resistencia a la ruptura promedio observada es $\bar{y} = 214$ psi. El valor del estadístico de prueba es

$$Z_0 = \frac{\bar{y} - \mu_0}{\sigma / \sqrt{n}} = \frac{214 - 200}{10 / \sqrt{4}} = 2.80$$

Si se especifica un error tipo I de $\alpha = 0.05$, en la tabla I del apéndice se encuentra que $Z_\alpha = Z_{0.05} = 1.645$. Por lo tanto, H_0 se rechaza y se concluye que la resistencia a la ruptura promedio del lote excede 200 psi.

....

Cuando no se conoce la varianza de la población, es necesario establecer el supuesto adicional de que la población sigue una distribución normal, aunque las desviaciones moderadas de la normalidad no afectarán seriamente los resultados.

Para probar H_0:$\mu = \mu_0$ en el caso de la varianza desconocida, se usa la varianza muestral S^2 para estimar σ^2. Al sustituir σ con S en la ecuación 2-35, se obtiene el estadístico de prueba

$$t_0 = \frac{\bar{y} - \mu_0}{S / \sqrt{n}} \tag{2-37}$$

La hipótesis nula $H_0: \mu = \mu_0$ se rechazaría si $|t_0| > t_{\alpha/2, n-1}$, donde $t_{\alpha/2, n-1}$ denota el punto porcentual $\alpha/2$ superior de la distribución t con $n - 1$ grados de libertad. El intervalo de confianza de $100(1 - \alpha)$ por ciento es en este caso

$$\bar{y} - t_{\alpha/2, n-1} S / \sqrt{n} \leq \mu \leq \bar{y} + t_{\alpha/2, n-1} S / \sqrt{n} \tag{2-38}$$

2-4.7 Resumen

En las tablas 2-3 y 2-4 se resumen los procedimientos de prueba estudiados aquí para las medias muestrales. Se muestran las regiones críticas para hipótesis alternativa tanto de una como de dos colas.

Tabla 2-3 Pruebas para medias con varianza conocida

Hipótesis	Estadístico de prueba	Criterios de rechazo
H_0:$\mu = \mu_0$ H_1:$\mu \neq \mu_0$		$\|Z_0\| > Z_{\alpha/2}$
H_0:$\mu = \mu_0$ H_1:$\mu < \mu_0$	$Z_0 = \frac{\bar{y} - \mu_0}{\sigma / \sqrt{n}}$	$Z_0 < -Z_\alpha$
H_0:$\mu = \mu_0$ H_1:$\mu > \mu_0$		$Z_0 > Z_\alpha$
H_0:$\mu_1 = \mu_2$ H_1:$\mu_2 \neq \mu_2$		$\|Z_0\| > Z_{\alpha/2}$
H_0:$\mu_1 = \mu_2$ H_1:$\mu_1 < \mu_2$	$Z_0 = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$	$Z_0 < -Z_\alpha$
H_0:$\mu_1 = \mu_2$ H_1:$\mu_1 > \mu_2$		$Z_0 > Z_\alpha$

Tabla 2-4 Pruebas para medias de distribuciones normales, varianza desconocida

Hipótesis	Estadístico de prueba	Criterios de rechazo
$H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$		$\lvert t_0 \rvert > t_{\alpha/2,n-1}$
$H_0: \mu = \mu_0$ $H_1: \mu < \mu_0$	$t_0 = \dfrac{\bar{y} - \mu_0}{S/\sqrt{n}}$	$t_0 < -t_{\alpha,n-1}$
$H_0: \mu = \mu_0$ $H_1: \mu > \mu_0$		$t_0 > t_{\alpha,n-1}$
	si $\sigma_1^2 = \sigma_2^2$	
$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 \neq \mu_2$	$t_0 = \dfrac{\bar{y}_1 - \bar{y}_2}{S_p\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$ $v = n_1 + n_2 - 2$	$\lvert t_0 \rvert > t_{\alpha/2,v}$
	si $\sigma_1^2 \neq \sigma_2^2$	
$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 < \mu_2$	$t_0 = \dfrac{\bar{y}_1 - \bar{y}_2}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}}$	$t_0 < -t_{\alpha,v}$
$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 > \mu_2$	$v = \dfrac{\left(\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}\right)^2}{\dfrac{(S_1^2/n_1)^2}{n_1 - 1} + \dfrac{(S_2^2/n_2)^2}{n_2 - 1}}$	$t_0 > t_{\alpha,v}$

2-5 INFERENCIAS ACERCA DE LAS DIFERENCIAS EN LAS MEDIAS, DISEÑOS DE COMPARACIONES PAREADAS

2-5.1 El problema de las comparaciones pareadas

En algunos experimentos comparativos simples puede conseguirse un mejoramiento significativo de la precisión haciendo comparaciones de observaciones pareadas del material experimental. Por ejemplo, considere una máquina para probar la dureza que presiona una barra con una punta afilada sobre un ejemplar de prueba de metal con una fuerza conocida. Al medir la profundidad de la depresión producida por la punta, se determina la dureza del ejemplar de prueba. En esta máquina pueden instalarse dos puntas diferentes y aun cuando la precisión (la variabilidad) de las mediciones hechas con las dos puntas parece ser la misma, se sospecha que una de las puntas produce diferentes lecturas de la dureza que la otra.

Sería posible realizar un experimento de la siguiente manera. Podrían seleccionarse al azar varios ejemplares de prueba del metal (por ejemplo, 20). La mitad de estos ejemplares de prueba podrían probarse con la punta 1 y la otra mitad con la punta 2. La asignación exacta de los ejemplares a las puntas se determinaría de manera aleatoria. Puesto que se trata de un diseño completamente aleatorizado, la dureza promedio de las dos muestras podría compararse utilizando la prueba *t* descrita en la sección 2-4.

Al reflexionar un poco al respecto, se descubriría una seria desventaja del diseño completamente aleatorizado en este problema. Suponga que los ejemplares de prueba del metal se cortaron de barras di-

ferentes que se fabricaron a temperaturas diferentes o que no fueran exactamente homogéneos en cualquier otra forma que pudiera afectar la dureza. Esta falta de homogeneidad entre los ejemplares contribuirá a la variabilidad de las mediciones de la dureza y tenderá a inflar el error experimental, haciendo más difícil detectar una diferencia real entre las puntas.

Para protegerse de esta posibilidad, considere un diseño experimental alternativo. Suponga que cada ejemplar de prueba tiene el tamaño suficiente para que puedan hacerse en él *dos* determinaciones de la dureza. Este diseño alternativo consistiría en dividir cada ejemplar de prueba en dos secciones, para después asignar de manera aleatoria una punta a una mitad de cada ejemplar de prueba y la otra punta a la otra mitad. El orden en que se prueban las puntas en un ejemplar de prueba particular se seleccionaría al azar. El experimento, cuando se llevó a cabo de acuerdo con este diseño con 10 ejemplares de prueba, produjo los datos (codificados) que se muestran en la tabla 2-5.

Un **modelo estadístico** que describe los datos de este experimento puede expresarse como

$$y_{ij} = \mu_i + \beta_j + \varepsilon_{ij} \begin{cases} i = 1, 2 \\ j = 1, 2, \ldots, 10 \end{cases} \tag{2-39}$$

donde y_{ij} es la observación de la dureza para la punta i en el ejemplar de prueba j, μ_i es la verdadera dureza promedio de la punta i-ésima, β_j es un efecto sobre la dureza debido al ejemplar de prueba j-ésimo, y ε_{ij} es el error experimental aleatorio con media cero y varianza σ_i^2. Es decir, σ_1^2 es la varianza de las mediciones de la dureza hechas con la punta 1 y σ_2^2 es la varianza de las mediciones de la dureza hechas con la punta 2.

Observe que si se calcula la diferencia pareada j-ésima

$$d_j = y_{1j} - y_{2j} \qquad j = 1, 2, \ldots, 10 \tag{2-40}$$

el valor esperado de esta diferencia es

$$\begin{aligned} \mu_d &= E(d_j) \\ &= E(y_{1j} - y_{2j}) \\ &= E(y_{1j}) - E(y_{2j}) \\ &= \mu_1 + \beta_j - (\mu_2 + \beta_j) \\ &= \mu_1 - \mu_2 \end{aligned}$$

Es decir, pueden hacerse inferencias acerca de la diferencia en las lecturas de la dureza promedio de las dos puntas $\mu_1 - \mu_2$ haciendo inferencias acerca de la media de las diferencias μ_d. Observe que el efecto adi-

Tabla 2-5 Datos del experimento de la prueba de la dureza

Ejemplar de prueba	Punta 1	Punta 2
1	7	6
2	3	3
3	3	5
4	4	3
5	8	8
6	3	2
7	2	4
8	9	9
9	5	4
10	4	5

tivo de las β_j de los ejemplares de prueba se cancela cuando las observaciones están pareadas de esta manera.

Probar H_0:$\mu_1 = \mu_2$ es equivalente a probar

$$H_0: \mu_d = 0$$
$$H_1: \mu_d \neq 0$$

El estadístico de prueba para esta hipótesis es

$$t_0 = \frac{\bar{d}}{S_d / \sqrt{n}} \tag{2-41}$$

donde

$$\bar{d} = \frac{1}{n}\sum_{j=1}^{n} d_j \tag{2-42}$$

es la media muestral de las diferencias y

$$S_d = \left[\frac{\sum_{j=1}^{n} (d_j - \bar{d})^2}{n-1}\right]^{1/2} = \left[\frac{\sum_{j=1}^{n} d_j^2 - \frac{1}{n}\left(\sum_{j=1}^{n} d_j\right)^2}{n-1}\right]^{1/2} \tag{2-43}$$

es la desviación estándar muestral de las diferencias. H_0: $\mu_d = 0$ se rechazaría si $|t_0| > t_{\alpha/2, n-1}$. Debido a que las observaciones de los niveles del factor están "pareadas" en cada unidad experimental, a este procedimiento suele llamársele **prueba *t* pareada**.

Por los datos de la tabla 2-5, se encuentra

$$d_1 = 7-6 = 1 \qquad d_6 = 3-2 = 1$$
$$d_2 = 3-3 = 0 \qquad d_7 = 2-4 = -2$$
$$d_3 = 3-5 = -2 \qquad d_8 = 9-9 = 0$$
$$d_4 = 4-3 = 1 \qquad d_9 = 5-4 = 1$$
$$d_5 = 8-8 = 0 \qquad d_{10} = 4-5 = -1$$

Por lo tanto,

$$\bar{d} = \frac{1}{n}\sum_{j=1}^{n} d_j = \frac{1}{10}(-1) = -0.10$$

$$S_d = \left[\frac{\sum_{j=1}^{n} d_j^2 - \frac{1}{n}\left(\sum_{j=1}^{n} d_j\right)^2}{n-1}\right]^{1/2} = \left[\frac{13 - \frac{1}{10}(-1)^2}{10-1}\right]^{1/2} = 1.20$$

Suponga que se elige $\alpha = 0.05$. Entonces, para tomar una decisión se calcularía t_0 y H_0 se rechazaría si $|t_0| > t_{0.025, 9} = 2.262$.

El valor calculado del estadístico de prueba *t* pareada es

$$t_0 = \frac{\bar{d}}{S_d / \sqrt{n}} = \frac{-0.10}{1.20/\sqrt{10}} = -0.26$$

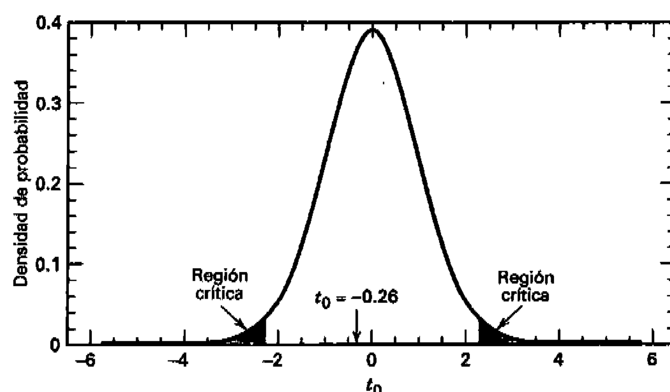


Figura 2-13 La distribución de referencia (t con 9 grados de libertad) para el problema de la prueba de la dureza.

y como $|t_0| = 0.26 \ngtr t_{0.025,9} = 2.262$, la hipótesis H_0:$\mu_d = 0$ no puede rechazarse. Es decir, no hay evidencia que indique que las dos puntas producen lecturas de la dureza diferentes. En la figura 2-13 se muestra la distribución de t_0 con 9 grados de libertad, la distribución de referencia para esta prueba, con el valor de t_0 indicado en relación con la región crítica.

En la tabla 2-6 se muestra la salida de computadora del procedimiento para la prueba t pareada de Minitab para este problema. Observe que el valor P para esta prueba es $P \simeq 0.80$, lo cual implica que no puede rechazarse la hipótesis nula con *ningún* nivel de significación razonable.

2-5.2 Ventajas del diseño de comparaciones pareadas

Al diseño que se utilizó en realidad para este experimento se le llama **diseño de comparaciones pareadas**, el cual ilustra el principio de la formación de bloques comentado en la sección 1-3. De hecho, es un caso especial de un tipo de diseño más general llamado **diseño de bloques aleatorizados**. El término *bloque* se refiere a una unidad experimental relativamente homogénea (en el caso tratado aquí, los ejemplares de prueba del metal son los bloques), y el bloque representa una restricción sobre la aleatorización completa debido a que las combinaciones de los tratamientos sólo se aleatorizan dentro del bloque. En el capítulo 4 se examina este tipo de diseños. En dicho capítulo el modelo matemático del diseño, la ecuación 2-39, se escribe en una forma ligeramente diferente.

Tabla 2-6 Resultados de Minitab de la prueba t pareada para el ejemplo de la prueba de la dureza

Prueba t pareada e intervalo de confianza

```
Paired T for Tip 1 - Tip 2

                 N        Mean       StDev     SE Mean
Tip 1           10       4.800       2.394       0.757
Tip 2           10       4.900       2.234       0.706
Difference      10      -0.100       1.197       0.379

95% CI for mean difference: (-0.956, 0.756)
t-Test of mean difference = 0 (vs not = 0):
T-Value = -0.26  P-Value = 0.798
```

Antes de dejar este experimento, es necesario destacar varios puntos. Observe que, aun cuando se han hecho $2n = 2(10) = 20$ observaciones, se cuenta únicamente con $n - 1 = 9$ grados de libertad para el estadístico t. (Se sabe que conforme se incrementan los grados de libertad para t, la prueba se hace más sensible.) Al hacer la formación de bloques o pareo, se han "perdido" en realidad $n - 1$ grados de libertad, pero se espera haber ganado un mejor conocimiento de la situación al eliminar una fuente adicional de variabilidad (la diferencia entre los ejemplares de prueba). Puede obtenerse una indicación de la calidad de la información producida por el diseño pareado comparando la desviación estándar S_d de las diferencias con la desviación estándar combinada S_p que habría resultado si el experimento se hubiera conducido de manera completamente aleatorizada y se hubieran obtenido los datos de la tabla 2-5. Al utilizar los datos de la tabla 2-5 como dos muestras independientes, la desviación estándar combinada que se calcula con la ecuación 2-25 es $S_p = 2.32$. Al comparar este valor con $S_d = 1.20$, se observa que la formación de bloques o pareo ha reducido la estimación de la variabilidad en cerca de 50%. Esta información también puede expresarse en términos de un intervalo de confianza para $\mu_1 - \mu_2$. Utilizando los datos pareados, un intervalo de confianza de 95% para $\mu_1 - \mu_2$ es

$$\bar{d} \pm t_{0.025,9} S_d / \sqrt{n}$$

$$-0.10 \pm (2.262)(1.20) / \sqrt{10}$$
$$-0.10 \pm 0.86$$

Recíprocamente, al utilizar el análisis combinado o independiente, un intervalo de confianza de 95% para $\mu_1 - \mu_2$ es

$$\bar{y}_1 - \bar{y}_2 \pm t_{0.025,18} S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$4.80 - 4.90 \pm (2.101)(2.32)\sqrt{\tfrac{1}{10} + \tfrac{1}{10}}$$
$$-0.10 \pm 2.18$$

El intervalo de confianza basado en el análisis pareado tiene una anchura sensiblemente menor que el intervalo de confianza del análisis independiente. Esto ilustra la propiedad de **reducción del ruido** de la formación de bloques.

La formación de bloques no es siempre la mejor estrategia de diseño. Si la variabilidad dentro de los bloques es la misma que la variabilidad entre los bloques, la varianza de $\bar{y}_1 - \bar{y}_2$ será la misma independientemente del diseño que se use. De hecho, la formación de bloques en esta situación sería una elección de diseño pobre porque la formación de bloques produce la pérdida de $n - 1$ grados de libertad y llevará en realidad a un intervalo de confianza con una anchura mayor para $\mu_1 - \mu_2$. En el capítulo 4 se ofrece una revisión más amplia de la formación de bloques.

2-6 INFERENCIAS ACERCA DE LAS VARIANZAS DE DISTRIBUCIONES NORMALES

En muchos experimentos, el interés se encuentra en las posibles diferencias en la respuesta media de dos tratamientos. Sin embargo, en algunos experimentos es la comparación de la variabilidad en los datos lo que es importante. En la industria de alimentos y bebidas, por ejemplo, es importante que la variabilidad del equipo de llenado sea pequeña para que todos los empaques estén cerca del peso neto nominal o el

volumen del contenido neto nominal. En los laboratorios químicos, tal vez quiera compararse la variabilidad de dos métodos de análisis. A continuación se examinan brevemente las pruebas de hipótesis y los intervalos de confianza para las varianzas de distribuciones normales. A diferencia de las pruebas para las medias, los procedimientos para las pruebas de varianzas son bastante más sensibles al supuesto de normalidad. En el apéndice 2A de Davies [36] hay un buen análisis del supuesto de normalidad.

Suponga que quiere probarse la hipótesis de que la varianza de una población normal es igual a una constante, por ejemplo, σ_0^2. Expresado en términos formales, quiere probarse

$$\begin{aligned} H_0&: \sigma^2 = \sigma_0^2 \\ H_1&: \sigma^2 \neq \sigma_0^2 \end{aligned} \tag{2-44}$$

El estadístico de prueba para la ecuación 2-44 es

$$\chi_0^2 = \frac{SS}{\sigma_0^2} = \frac{(n-1)S^2}{\sigma_0^2} \tag{2-45}$$

donde $SS = \sum_{i=1}^{n}(y_i - y)^2$ es la suma de cuadrados corregida de las observaciones muestrales. La distribución de referencia apropiada para χ_0^2 es la distribución ji-cuadrada con $n-1$ grados de libertad. La hipótesis nula se rechaza si $\chi_0^2 > \chi_{\alpha/2,n-1}^2$ o si $\chi_0^2 < \chi_{1-(\alpha/2),n-1}^2$, donde $\chi_{\alpha/2,n-1}^2$ y $\chi_{1-(\alpha/2),n-1}^2$ son los puntos porcentuales $\alpha/2$ superior y $1-(\alpha/2)$ inferior de la distribución ji-cuadrada con $n-1$ grados de libertad, respectivamente. En la tabla 2-7 se presentan las regiones críticas para las hipótesis alternativas de una cola. El intervalo de confianza de $100(1-\alpha)$ por ciento para σ^2 es

$$\frac{(n-1)S^2}{\chi_{\alpha/2,n-1}^2} \leq \sigma^2 \leq \frac{(n-1)S^2}{\chi_{1-(\alpha/2),n-1}^2} \tag{2-46}$$

Considere ahora la prueba de la igualdad de las varianzas de dos poblaciones normales. Si se toman muestras aleatorias independientes de tamaño n_1 y n_2 de las poblaciones 1 y 2, respectivamente, el estadístico de prueba para

$$\begin{aligned} H_0&: \sigma_1^2 = \sigma_2^2 \\ H_1&: \sigma_1^2 \neq \sigma_2^2 \end{aligned} \tag{2-47}$$

es el cociente de las varianzas muestrales

$$F_0 = \frac{S_1^2}{S_2^2} \tag{2-48}$$

La distribución de referencia apropiada para F_0 es la distribución F con $n_1 - 1$ grados de libertad en el numerador y $n_2 - 1$ grados de libertad en el denominador. La hipótesis nula se rechazaría si $F_0 > F_{\alpha/2,n_1-1,n_2-1}$ o si $F_0 < F_{1-(\alpha/2),n_1-1,n_2-1}$, donde $F_{\alpha/2,n_1-1,n_2-1}$ y $F_{1-(\alpha/2),n_1-1,n_2-1}$ denotan los puntos porcentuales $\alpha/2$ superior y $1-(\alpha/2)$ inferior de la distribución F con $n_1 - 1$ y $n_2 - 1$ grados de libertad. En la tabla IV del apéndice sólo aparecen los puntos porcentuales para la cola superior de F; sin embargo, los puntos de las colas superior e inferior se relacionan por

$$F_{1-\alpha,v_1,v_2} = \frac{1}{F_{\alpha,v_2,v_1}} \tag{2-49}$$

Tabla 2-7 Pruebas para las varianzas de distribuciones normales

Hipótesis	Estadístico de prueba	Criterios de rechazo
$H_0:\sigma^2 = \sigma_0^2$ $H_1:\sigma^2 \neq \sigma_0^2$		$\chi_0^2 > \chi^2_{\alpha/2,n-1}$ o $\chi_0^2 < \chi^2_{1-\alpha/2,n-1}$
$H_0:\sigma^2 = \sigma_0^2$ $H_1:\sigma^2 < \sigma_0^2$	$\chi_0^2 = \dfrac{(n-1)S^2}{\sigma_0^2}$	$\chi_0^2 < \chi^2_{1-\alpha,n-1}$
$H_0:\sigma^2 = \sigma_0^2$ $H_1:\sigma^2 > \sigma_0^2$		$\chi_0^2 > \chi^2_{\alpha,n-1}$
$H_0:\sigma_1^2 = \sigma_2^2$ $H_1:\sigma_1^2 \neq \sigma_2^2$	$F_0 = \dfrac{S_1^2}{S_2^2}$	$F_0 > F_{\alpha/2,n_1-1,n_2-1}$ o $F_0 < F_{1-\alpha/2,n_1-1,n_2-1}$
$H_0:\sigma_1^2 = \sigma_2^2$ $H_1:\sigma_1^2 < \sigma_2^2$	$F_0 = \dfrac{S_2^2}{S_1^2}$	$F_0 > F_{\alpha,n_2-1,n_1-1}$
$H_0:\sigma_1^2 = \sigma_2^2$ $H_1:\sigma_1^2 > \sigma_2^2$	$F_0 = \dfrac{S_1^2}{S_2^2}$	$F_0 > F_{\alpha,n_1-1,n_2-1}$

En el capítulo 3, sección 3-4.3, se analizan los procedimientos de prueba para más de dos varianzas. Se revisará también el uso de la varianza o la desviación estándar como variable de respuesta en situaciones experimentales más generales.

EJEMPLO 2-2 ....

Un ingeniero químico investiga la variabilidad inherente de dos tipos de equipo de prueba que pueden usarse para monitorear la producción de un proceso. El ingeniero sospecha que el equipo antiguo, tipo 1, tiene una varianza mayor que la del equipo nuevo. Por lo tanto, quiere probar las hipótesis

$$H_0: \sigma_1^2 = \sigma_2^2$$
$$H_1: \sigma_1^2 > \sigma_2^2$$

Se toman dos muestras aleatorias de $n_1 = 12$ y $n_2 = 10$ observaciones, y las varianzas muestrales son $S_1^2 = 14.5$ y $S_2^2 = 10.8$. El estadístico de prueba es

$$F_0 = \frac{S_1^2}{S_2^2} = \frac{14.5}{10.8} = 1.34$$

En la tabla IV del apéndice se encuentra que $F_{0.05,\,11,\,9} = 3.10$, por lo que no puede rechazarse la hipótesis nula. Es decir, se ha encontrado evidencia estadística insuficiente para concluir que la varianza del equipo antiguo sea mayor que la varianza del equipo nuevo.

....

El intervalo de confianza de $100(1 - \alpha)$ por ciento para el cociente de las varianzas poblacionales σ_1^2 / σ_2^2 es

$$\frac{S_1^2}{S_2^2} F_{1-\alpha/2,n_2-1,n_1-1} \leq \frac{\sigma_1^2}{\sigma_2^2} \leq \frac{S_1^2}{S_2^2} F_{\alpha/2,n_2-1,n_1-1} \tag{2-50}$$

Para ilustrar el uso de la ecuación 2-50, el intervalo de confianza de 95% para el cociente de las varianzas σ_1^2 / σ_2^2 del ejemplo 2-2 es, utilizando $F_{0.025,\,9,11} = 3.59$ y $F_{0.975,9,11} = 1/F_{0.025,11,9} = 1/3.92 = 0.255$,

$$\frac{14.5}{10.8}(0.255) \leq \frac{\sigma_1^2}{\sigma_2^2} \leq \frac{14.5}{10.8}(3.59)$$

$$0.34 \leq \frac{\sigma_1^2}{\sigma_2^2} \leq 4.81$$

2-7 PROBLEMAS

2-1. Se requiere que la resistencia a la ruptura de una fibra sea de por lo menos 150 psi. La experiencia pasada indica que la desviación estándar de la resistencia a la ruptura es $\sigma = 3$ psi. Se prueba una muestra aleatoria de cuatro ejemplares de prueba, y los resultados son $y_1 = 145$, $y_2 = 153$, $y_3 = 150$ y $y_4 = 147$.

a) Enunciar las hipótesis que el lector considere que deberían probarse en este experimento.
b) Probar estas hipótesis utilizando $\alpha = 0.05$. ¿A qué conclusiones se llega?
c) Encontrar el valor *P* para la prueba del inciso *b*.
d) Construir un intervalo de confianza de 95% para la resistencia a la ruptura promedio.

2-2. Supuestamente, la viscosidad de un detergente líquido debe promediar 800 centistokes a 25°C. Se colecta una muestra aleatoria de 16 lotes del detergente, y la viscosidad promedio es 812. Suponga que se sabe que la desviación estándar de la viscosidad es $\sigma = 25$ centistokes.

a) Enunciar las hipótesis que deberán probarse.
b) Probar estas hipótesis utilizando $\alpha = 0.05$. ¿A qué conclusiones se llega?
c) ¿Cuál es el valor *P* para la prueba?
d) Encontrar un intervalo de confianza de 95% para la media.

2-3. Los diámetros de las flechas de acero producidas en cierto proceso de manufactura deberán tener un promedio de 0.255 pulgadas. Se sabe que el diámetro tiene una desviación estándar de $\sigma = 0.0001$ pulgadas. Una muestra aleatoria de 10 flechas tiene un diámetro promedio de 0.2545 pulgadas.

a) Establecer las hipótesis apropiadas para la media μ.
b) Probar estas hipótesis utilizando $\alpha = 0.05$. ¿A qué conclusiones se llega?
c) Encontrar el valor *P* para esta prueba.
d) Construir un intervalo de confianza de 95% para el diámetro promedio de las flechas.

2-4. Una variable aleatoria con una distribución normal tiene una media desconocida μ y varianza $\sigma^2 = 9$. Encontrar el tamaño de la muestra que se necesita para construir un intervalo de confianza de 95% para la media, cuya anchura total sea de 1.0.

2-5. La vida de anaquel de una bebida carbonatada es motivo de interés. Se seleccionan 10 botellas al azar y se prueban, obteniéndose los siguientes resultados:

Días	
108	138
124	163
124	159
106	134
115	139

a) Quiere demostrarse que la vida media de anaquel excede los 120 días. Establecer las hipótesis apropiadas para investigar esta afirmación.
b) Probar estas hipótesis utilizando $\alpha = 0.01$. ¿A qué conclusiones se llega?

c) Encontrar el valor *P* para la prueba del inciso *b*.

d) Construir un intervalo de confianza de 99% para la vida media de anaquel.

2-6. Considere los datos de la vida de anaquel del problema 2-5. ¿La vida de anaquel puede describirse o modelarse adecuadamente con una distribución normal? ¿Qué efecto tendría la violación de este supuesto sobre el procedimiento de prueba usado para resolver el problema 2-5?

2-7. El tiempo para reparar un instrumento electrónico es una variable aleatoria medida en horas que sigue una distribución normal. El tiempo de reparación de 16 de estos instrumentos elegidos al azar es el siguiente:

Horas			
159	280	101	212
224	379	179	264
222	362	168	250
149	260	485	170

a) Quiere saberse si el tiempo de reparación promedio excede 225 horas. Establecer las hipótesis apropiadas para investigar esta cuestión.

b) Probar las hipótesis que se formularon en el inciso *a*. ¿A qué conclusiones se llega? Utilizar $\alpha = 0.05$.

c) Encontrar el valor *P* para la prueba.

d) Construir un intervalo de confianza de 95% para el tiempo de reparación promedio.

2-8. Considere nuevamente los datos del tiempo de reparación del problema 2-7. En opinión del lector, ¿el tiempo de reparación puede modelarse de manera adecuada con una distribución normal?

2-9. Se utilizan dos máquinas para llenar botellas de plástico con un volumen neto de 16.0 onzas. Puede suponerse que el proceso de llenado es normal, con desviaciones estándar de $\sigma_1 = 0.015$ y $\sigma_2 = 0.018$. El departamento de ingeniería de calidad sospecha que ambas máquinas llenan el mismo volumen neto, sin importar si este volumen es 16.0 onzas o no. Se realiza un experimento tomando una muestra aleatoria de la producción de cada máquina.

Máquina 1		Máquina 2	
16.03	16.01	16.02	16.03
16.04	15.96	15.97	16.04
16.05	15.98	15.96	16.02
16.05	16.02	16.01	16.01
16.02	15.99	15.99	16.00

a) Enunciar las hipótesis que deberán probarse en este experimento.

b) Probar estas hipótesis utilizando $\alpha = 0.05$. ¿A qué conclusiones se llega?

c) Encontrar el valor *P* para esta prueba.

d) Encontrar un intervalo de confianza de 95% para la diferencia en el volumen de llenado promedio de las dos máquinas.

2-10. Un fabricante de calculadoras electrónicas puede usar dos tipos de plástico. La resistencia a la ruptura de este plástico es importante. Se sabe que $\sigma_1 = \sigma_2 = 1.0$ psi. De muestras aleatorias de $n_1 = 10$ y $n_2 = 12$ se obtiene $\bar{y}_1 = 162.5$ y $\bar{y}_2 = 155.0$. La compañía no empleará el plástico 1 a menos que su resistencia a la ruptura exceda la del plástico 2 por al menos 10 psi. Con base en la información muestral, ¿deberá usarse el plástico 1? Para responder esta pregunta se deben establecer y probar las hipótesis apropiadas utilizando $\alpha = 0.01$. Construir un intervalo de confianza de 99% para la verdadera diferencia media en la resistencia a la ruptura.

2-11. A continuación se presenta el tiempo de combustión de dos cohetes químicos con formulaciones diferentes. Los ingenieros de diseño se interesan tanto en la media como en la varianza del tiempo de combustión.

Tipo 1		Tipo 2	
65	82	64	56
81	67	71	69
57	59	83	74
66	75	59	82
82	70	65	79

a) Probar la hipótesis de que las dos varianzas son iguales. Utilizar $\alpha = 0.05$.
b) Utilizando los resultados del inciso *a*, probar la hipótesis de que los tiempos de combustión promedio son iguales. Utilizar $\alpha = 0.05$. ¿Cuál es el valor *P* para esta prueba?
c) Comentar el papel del supuesto de normalidad en este problema. Verificar el supuesto de normalidad para ambos tipos de cohetes.

2-12. En un artículo de *Solid State Technology*, "Diseño ortogonal para optimización de procesos y su aplicación en el grabado químico con plasma" de G.Z. Yin y D.W. Jillie, se describe un experimento para determinar el efecto de la velocidad del flujo de C_2F_6 sobre la uniformidad del grabado en una oblea de silicio usada en la fabricación de circuitos integrados. Los datos de la velocidad del flujo son los siguientes:

Flujo de C_2F_6	Observación de la uniformidad					
	1	2	3	4	5	6
125	2.7	4.6	2.6	3.0	3.2	3.8
200	4.6	3.4	2.9	3.5	4.1	5.1

a) ¿La velocidad del flujo de C_2F_6 afecta la uniformidad del grabado promedio? Utilizar $\alpha = 0.05$.
b) ¿Cuál es el valor *P* para la prueba del inciso *a*?
c) ¿La velocidad del flujo de C_2F_6 afecta la variabilidad de una oblea a otra en la uniformidad del grabado? Utilizar $\alpha = 0.05$.
d) Trazar diagramas de caja que ayuden a interpretar los datos de este experimento.

2-13. Se instala un nuevo dispositivo de filtrado en una unidad química. Antes de instalarlo, de una muestra aleatoria se obtuvo la siguiente información sobre el porcentaje de impurezas: $\bar{y}_1 = 12.5$, $S_1^2 = 101.17$ y $n_1 = 8$. Después de instalarlo, de una muestra aleatoria se obtuvo $\bar{y}_2 = 10.2$, $S_2^2 = 94.73$, $n_2 = 9$.
a) ¿Puede concluirse que las dos varianzas son iguales? Utilizar $\alpha = 0.05$.
b) ¿El dispositivo de filtrado ha reducido de manera significativa el porcentaje de impurezas? Utilizar $\alpha = 0.05$.

2-14. Se hacen 20 observaciones de la uniformidad del grabado en obleas de silicio durante un experimento de evaluación de un grabador de plasma. Los datos son los siguientes:

5.34	6.65	4.76	5.98	7.25
6.00	7.55	5.54	5.62	6.21
5.97	7.35	5.44	4.39	4.98
5.25	6.35	4.61	6.00	5.32

a) Construir una estimación con un intervalo de confianza de 95% de σ^2.
b) Probar la hipótesis de que $\sigma^2 = 1.0$. Utilizar $\alpha = 0.05$. ¿A qué conclusiones se llega?

c) Comentar el supuesto de normalidad y su papel en este problema.

d) Verificar la normalidad construyendo una gráfica de probabilidad normal. ¿A qué conclusiones se llega?

2-15. Doce inspectores midieron el diámetro de un cojinete de bolas, utilizando cada uno dos tipos diferentes de calibradores. Los resultados fueron

Inspector	Calibrador 1	Calibrador 2
1	0.265	0.264
2	0.265	0.265
3	0.266	0.264
4	0.267	0.266
5	0.267	0.267
6	0.265	0.268
7	0.267	0.264
8	0.267	0.265
9	0.265	0.265
10	0.268	0.267
11	0.268	0.268
12	0.265	0.269

a) ¿Existe una diferencia significativa entre las medias de la población de mediciones de las que se seleccionaron las dos muestras? Utilizar $\alpha = 0.05$.

b) Encontrar el valor *P* para la prueba del inciso *a*.

c) Construir un intervalo de confianza de 95% para la diferencia en las mediciones de los diámetros promedio para los dos tipos de calibradores.

2-16. En un artículo de *Journal of Strain Analysis* (vol. 18, no. 2) se comparan varios procedimientos para predecir la resistencia al corte de vigas de placas de acero. Los datos para nueve vigas en la forma del cociente de la carga predicha y la observada para dos de estos procedimientos, los métodos Karlsruhe y Lehigh, son los siguientes

Viga	Método Karlsruhe	Método Lehigh
S1/1	1.186	1.061
S2/1	1.151	0.992
S3/1	1.322	1.063
S4/1	1.339	1.062
S5/1	1.200	1.065
S2/1	1.402	1.178
S2/2	1.365	1.037
S2/3	1.537	1.086
S2/4	1.559	1.052

a) ¿Existe alguna evidencia que apoye la afirmación de que hay una diferencia en el desempeño promedio entre los dos métodos? Utilizar $\alpha = 0.05$.

b) ¿Cuál es el valor *P* para la prueba del inciso *a*?

c) Construir un intervalo de confianza de 95% para la diferencia en la carga promedio predicha y la observada.

d) Investigar el supuesto de normalidad en ambas muestras.

e) Investigar el supuesto de normalidad para la diferencia en los cocientes para los dos métodos.

f) Comentar el papel del supuesto de normalidad en la prueba *t* pareada.

2-17. Se estudia la temperatura de deflexión bajo carga de dos formulaciones diferentes de un tubo de plástico ABS. Dos muestras de 12 observaciones cada una, se preparan utilizando cada formulación y las temperaturas de deflexión (en °F) se presentan abajo:

Formulación 1			Formulación 2		
206	193	192	177	176	198
188	207	210	197	185	188
205	185	194	206	200	189
187	189	178	201	197	203

a) Construir las gráficas de probabilidad normal para ambas muestras. ¿Estas gráficas apoyan los supuestos de normalidad y de la igualdad de la varianza de ambas muestras?

b) ¿Los datos apoyan la afirmación de que la temperatura promedio de deflexión bajo carga de la formulación 1 excede la de la formulación 2? Utilizar $\alpha = 0.05$.

c) ¿Cuál es el valor *P* para la prueba del inciso *a*?

2-18. Referirse a los datos del problema 2-17. ¿Los datos apoyan la afirmación de que la temperatura promedio de deflexión bajo carga de la formulación 1 excede la de la formulación 2 en al menos 3°F?

2-19. En la fabricación de semiconductores es común el uso del grabado químico húmedo para eliminar el silicio de la parte posterior de las obleas antes de la metalización. La rapidez del grabado es una característica importante de este proceso. Se están evaluando dos soluciones de grabado diferentes. Se grabaron ocho obleas seleccionadas al azar en cada solución, y las cifras de la rapidez del grabado observada (en milésimas de pulgada/min) se muestran abajo

Solución 1		Solución 2	
9.9	10.6	10.2	10.6
9.4	10.3	10.0	10.2
10.0	9.3	10.7	10.4
10.3	9.8	10.5	10.3

a) ¿Los datos indican que la afirmación de que ambas soluciones tienen la misma rapidez de grabado promedio es verdadera? Utilizar $\alpha = 0.05$ y suponer la igualdad de las varianzas.

b) Encontrar un intervalo de confianza de 95% para la diferencia en la rapidez de grabado promedio.

c) Usar gráficas de probabilidad normal para investigar la adecuación de los supuestos de normalidad e igualdad de las varianzas.

2-20. Se están comparando dos populares analgésicos con base en la rapidez de absorción del cuerpo. Específicamente, se afirma que la tableta 1 se absorbe con el doble de rapidez que la tableta 2. Suponer que σ_1^2 y σ_2^2 se conocen. Desarrollar un estadístico de prueba para

$$H_0{:}2\mu_1 = \mu_2$$
$$H_1{:}2\mu_1 \neq \mu_2$$

2-21. Suponga que se está probando

$$H_0{:}\mu_1 = \mu_2$$
$$H_1{:}\mu_1 \neq \mu_2$$

donde σ_1^2 y σ_2^2 se conocen. Los recursos para hacer el muestreo son limitados, por lo que $n_1 + n_2 = N$. ¿Cómo deberán asignarse las *N* observaciones entre las dos poblaciones para obtener la prueba con la potencia más alta?

2-22. Desarrollar la ecuación 2-46 para un intervalo de confianza de $100(1-\alpha)$ por ciento para la varianza de una distribución normal.

2-23. Desarrollar la ecuación 2-50 para un intervalo de confianza de $100(1-\alpha)$ por ciento para el cociente σ_1^2 / σ_2^2, donde σ_1^2 y σ_2^2 son las varianzas de dos distribuciones normales.

2-24. Desarrollar una ecuación para encontrar un intervalo de confianza de $100(1-\alpha)$ por ciento para la diferencia en las medias de dos distribuciones normales donde $\sigma_1^2 \neq \sigma_2^2$. Aplicar la ecuación desarrollada a los datos del experimento del cemento portland, y encontrar un intervalo de confianza de 95%.

2-25. Construir un conjunto de datos para los que el estadístico de prueba t pareada sea muy grande, pero para el cual el estadístico de prueba t de dos muestras o combinada usual sea pequeño. En general, describir cómo se crearon los datos. ¿Le da esto al lector alguna idea respecto de cómo funciona la prueba t pareada?

3 Experimentos con un solo factor: el análisis de varianza

En el capítulo 2 se analizaron los métodos para comparar dos condiciones o tratamientos. Por ejemplo, el experimento de la fuerza de la tensión de adhesión del cemento portland incluyó dos formulaciones diferentes del mortero. Otra forma de describir este experimento es como un experimento con un solo factor, con dos niveles del factor, donde el factor es la formulación del mortero y los dos niveles son los dos métodos diferentes para hacer la formulación. Muchos experimentos de este tipo involucran más de dos niveles del factor. En este capítulo se presentan los métodos para el diseño y el análisis de los experimentos con un solo factor con a niveles del mismo (o a tratamientos). Se supondrá que el experimento se ha aleatorizado completamente.

3-1 UN EJEMPLO

Un ingeniero de desarrollo de productos tiene interés en investigar la resistencia a la tensión de una fibra sintética nueva que se usará para hacer tela de camisas para caballero. El ingeniero sabe por experiencia previa que la resistencia a la tensión se afecta por el peso porcentual del algodón utilizado en la mezcla de materiales de la fibra. Además, sospecha que al aumentar el contenido de algodón se incrementará la resistencia, al menos en un principio. Sabe asimismo que el contenido de algodón deberá variar entre 10 y 40 por ciento para que el producto final tenga otras características de calidad que se desean (como la capacidad de ser sometido a un tratamiento de planchado permanente). El ingeniero decide probar ejemplares en cinco niveles del peso porcentual del algodón: 15, 20, 25, 30 y 35 por ciento. También decide probar cinco ejemplares en cada nivel del contenido de algodón.

Se trata de un ejemplo de un experimento con un solo factor con $a = 5$ **niveles** del factor y $n = 5$ **réplicas.** Las 25 corridas deberán realizarse de manera aleatoria. Para ilustrar cómo puede aleatorizarse el orden de las corridas, suponga que las corridas se numeran de la siguiente manera:

Peso porcentual del algodón	Número de corrida experimental				
15	1	2	3	4	5
20	6	7	8	9	10
25	11	12	13	14	15
30	16	17	18	19	20
35	21	22	23	24	25

Ahora se selecciona un número aleatorio entre 1 y 25. Suponga que este número es 8. Entonces la observación número 8 (20% de algodón) se corre primero. Este proceso se repetiría hasta que las 25 observaciones tengan asignada una posición en la secuencia de prueba.[1] Muchos paquetes de software de computadora para ayudar a los experimentadores a seleccionar y construir un diseño, aleatorizan el orden de las corridas utilizando números aleatorios de esta manera.

Suponga que la secuencia de prueba obtenida es

Secuencia de prueba	Número de corrida	Peso porcentual del algodón
1	8	20
2	18	30
3	10	20
4	23	35
5	17	30
6	5	15
7	14	25
8	6	20
9	15	25
10	20	30
11	9	20
12	4	15
13	12	25
14	7	20
15	1	15
16	24	35
17	21	35
18	11	25
19	2	15
20	13	25
21	22	35
22	16	30
23	25	35
24	19	30
25	3	15

Esta secuencia de prueba aleatorizada es necesaria para evitar que los efectos de variables perturbadoras desconocidas —las cuales quizá varíen fuera de control durante el experimento— contaminen los resultados. Para ilustrar esto, suponga que las 25 corridas de prueba tuvieran que realizarse en el orden original no aleatorizado (es decir, primero se prueban los cinco ejemplares con 15% de algodón, después se prue-

[1] La única restricción sobre la aleatorización en este caso, es que si se saca de nuevo el mismo número (es decir, 8), se descarta. Se trata de una restricción secundaria y se ignora.

Tabla 3-1 Datos (en lb/pulgada²) del experimento de la resistencia a la tensión

Peso porcentual del algodón	Observaciones 1	2	3	4	5	Total	Promedio
15	7	7	15	11	9	49	9.8
20	12	17	12	18	18	77	15.4
25	14	18	18	19	19	88	17.6
30	19	25	22	19	23	108	21.6
35	7	10	11	15	11	54	10.8
						376	15.04

ban los cinco ejemplares con 20% de algodón, etc.). Si la máquina empleada para probar la resistencia a la tensión presenta un efecto de calentamiento tal que entre más tiempo esté funcionando sean menores las lecturas de la resistencia a la tensión observadas, el efecto del calentamiento contaminará potencialmente los datos de la resistencia a la tensión y destruirá la validez del experimento.

Suponga que el ingeniero corre la prueba en el orden aleatorio que se ha determinado. En la tabla 3-1 se muestran las observaciones que obtiene para la resistencia a la tensión.

Siempre es una buena idea examinar **gráficamente** los datos experimentales. En la figura 3-1 se muestran los **diagramas de caja** para la resistencia a la tensión con cada nivel del peso porcentual de algodón, y en la figura 3-2 se ilustra un **diagrama de dispersión** de la resistencia a la tensión contra el peso porcentual del algodón. En la figura 3-2, los puntos rellenos son las observaciones individuales y los círculos huecos son los promedios de la resistencia a la tensión observada. Ambas gráficas indican que la resistencia a la tensión se incrementa cuando el contenido de algodón se incrementa, hasta cerca de 30% de algodón. Después de 30% de algodón, hay un marcado descenso de la resistencia a la tensión. No hay evidencia sólida que sugiera que la variabilidad de la resistencia a la tensión alrededor del promedio dependa del peso porcentual del algodón. Con base en este análisis gráfico simple, se tienen firmes sospechas de que 1) el contenido de algodón afecta la resistencia a la tensión y 2) alrededor de 30% de algodón produce la resistencia máxima.

Suponga que se quiere ser más **objetivo** en el análisis de los datos. Específicamente, imagine que quieren probarse las diferencias entre las resistencias a la tensión promedio con todos los niveles $a = 5$ del

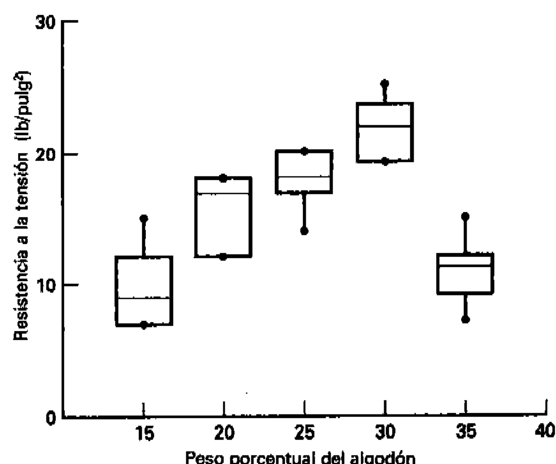


Figura 3-1 Diagramas de caja de la resistencia a la tensión contra el peso porcentual del algodón.

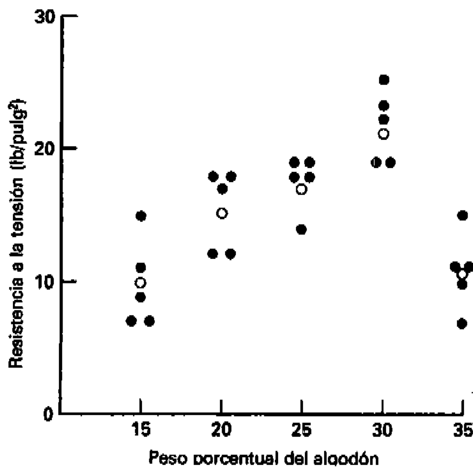


Figura 3-2 Diagrama de dispersión de la resistencia a la tensión contra el peso porcentual del algodón.

peso porcentual del algodón. Por lo tanto, el interés se centra en probar la igualdad de las cinco medias. Pudiera parecer que este problema se resolvería realizando una prueba t para todos los pares de medias posibles. Sin embargo, no es ésta la mejor solución de este problema, porque llevaría a una distorsión considerable en el error tipo I. Por ejemplo, suponga que quiere probarse la igualdad de las cinco medias usando comparaciones por pares. Hay 10 pares posibles, y si la probabilidad de aceptar correctamente la hipótesis nula en cada prueba individual es de $1 - \alpha = 0.95$, la probabilidad de aceptar correctamente la hipótesis nula en las 10 pruebas es de $(0.95)^{10} = 0.60$ si las pruebas son independientes. Por lo tanto, ha ocurrido un incremento sustancial en el error tipo I.

El procedimiento correcto para probar la igualdad de varias medias es el **análisis de varianza.** Sin embargo, el análisis de varianza tiene un rango de aplicaciones mucho más amplio que el problema anterior. Probablemente sea la técnica más útil en el campo de la inferencia estadística.

3-2 EL ANÁLISIS DE VARIANZA

Suponga que se tienen a **tratamientos** o **niveles** diferentes de **un solo factor** que quieren compararse. La respuesta observada de cada uno de los a tratamientos es una variable aleatoria. Los datos aparecerían como en la tabla 3-2. Una entrada de la tabla 3-2 (por ejemplo, y_{ij}) representa la observación j-ésima to-

Tabla 3-2 Datos típicos de un experimento de un solo factor

Tratamiento (nivel)	Observaciones				Totales	Promedios
1	y_{11}	y_{12}	$\cdots$	y_{1n}	$y_{1.}$	$\bar{y}_{1.}$
2	y_{21}	y_{22}	$\cdots$	y_{2n}	$y_{2.}$	$\bar{y}_{2.}$
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$	$\vdots$
a	y_{a1}	y_{a2}	$\cdots$	y_{an}	$y_{a.}$	$\bar{y}_{a.}$
					$y_{..}$	$\bar{y}_{..}$

mada bajo el nivel del factor o tratamiento i. Habrá, en general, n observaciones bajo el tratamiento i-ésimo. Observe que la tabla 3-2 es el caso general de los datos del experimento de la resistencia a la tensión de la tabla 3-1.

Modelos para los datos

Se encontrará útil describir las observaciones de un experimento con un modelo. Una manera de escribir este modelo es

$$y_{ij} = \mu_i + \varepsilon_{ij} \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, n \end{cases} \tag{3-1}$$

donde y_{ij} es la observación ij-ésima, μ_i es la media del nivel del factor o tratamiento i-ésimo, y ε_{ij} es un componente del **error aleatorio** que incorpora todas las demás fuentes de variabilidad del experimento, incluyendo las mediciones, la variabilidad que surge de factores no controlados, las diferencias entre las unidades experimentales (como los materiales de prueba, etc.) a las que se aplican los tratamientos, y el ruido de fondo general en el proceso (ya sean la variabilidad con el tiempo, los efectos de variables ambientales, etc.). Es conveniente considerar que los errores tienen media cero, de tal modo que $E(y_{ij}) = \mu_i$.

A la ecuación 3-1 se le llama el **modelo de las medias.** Una forma alternativa de escribir un modelo de los datos es definiendo

$$\mu_i = \mu + \tau_i, \qquad i = 1, 2, \ldots, a$$

de tal modo que la ecuación 3-1 se convierte en

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij} \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, n \end{cases} \tag{3-2}$$

En esta forma del modelo, μ es un parámetro común a todos los tratamientos al que se llama la **media global**, y τ_i es un parámetro único del tratamiento i-ésimo al que se le llama el **efecto del tratamiento i-ésimo.** A la ecuación 3-2 se le llama por lo general el **modelo de los efectos.**

Tanto el modelo de las medias como el de los efectos son **modelos estadísticos lineales;** es decir, la variable de respuesta y_{ij} es una función lineal de los parámetros del modelo. Aun cuando ambas formas del modelo son útiles, el modelo de los efectos se encuentra con mayor frecuencia en la literatura del diseño experimental. Tiene cierto atractivo intuitivo por cuanto μ es una constante y los efectos de los tratamientos τ_i representan desviaciones de esta constante cuando se aplican los tratamientos específicos.

A la ecuación 3-2 (o a la 3-1) se le llama también el modelo del **análisis de varianza simple** o **de un solo factor (o dirección)**, porque únicamente se investiga un factor. Además, será un requisito que el experimento se lleve a cabo en orden aleatorio para que el ambiente en el que se apliquen los tratamientos (llamados con frecuencia **unidades experimentales**) sea lo más uniforme posible. Por lo tanto, el diseño experimental es un **diseño completamente aleatorizado.** Los objetivos serán probar las hipótesis apropiadas acerca de las medias de los tratamientos y estimarlas. Para probar las hipótesis, se supone que los errores del modelo son variables aleatorias que siguen una distribución normal e independiente con media cero y varianza σ^2. Se supone asimismo que la varianza σ^2 es constante para todos los niveles del factor. Esto implica que las observaciones

$$y_{ij} \sim N(\mu + \tau_i, \sigma^2)$$

y que las observaciones son mutuamente independientes.

¿Factor fijo o aleatorio?

El modelo estadístico (ecuación 3-2) describe dos situaciones diferentes con respecto a los efectos de los tratamientos. Primera, los a tratamientos pudieron ser elegidos expresamente por el experimentador. En esta situación quieren probarse hipótesis acerca de las medias de los tratamientos, y las conclusiones se aplicarán únicamente a los niveles del factor considerados en el análisis. Las conclusiones no pueden extenderse a tratamientos similares que no fueron considerados explícitamente. También se podría querer estimar los parámetros del modelo (μ, τ_i, σ^2). A éste se le llama el **modelo con efectos fijos.** De manera alternativa, los a tratamientos podrían ser una **muestra aleatoria** de una población más grande de tratamientos. En esta situación sería deseable poder extender las conclusiones (las cuales se basan en la muestra de los tratamientos) a la totalidad de los tratamientos de la población, sea que se hayan considerado explícitamente en el análisis o no. Aquí las τ_i son variables aleatorias, y el conocimiento de las τ_i particulares que se investigaron es relativamente inútil. Más bien, se prueban hipótesis acerca de la variabilidad de las τ_i y se intenta estimar su variabilidad. A éste se le llama el **modelo con efectos aleatorios** o **modelo de los componentes de la varianza.** La revisión de experimentos con factores aleatorios se pospondrá hasta el capítulo 12.

3-3 ANÁLISIS DEL MODELO CON EFECTOS FIJOS

En esta sección se desarrolla el análisis de varianza de un solo factor para el modelo con efectos fijos. Recuerde que $y_{i.}$ representa el total de las observaciones bajo el tratamiento i-ésimo. Sea que $\bar{y}_{i.}$ represente el promedio de las observaciones bajo el tratamiento i-ésimo. De manera similar, sea que $y_{..}$ represente el gran total de todas las observaciones y que $\bar{y}_{..}$ represente el gran promedio de todas las observaciones. Expresado simbólicamente,

$$\begin{aligned} y_{i.} &= \sum_{j=1}^{n} y_{ij}, \qquad \bar{y}_{i.} = y_{i.}/n \qquad i = 1, 2, \ldots, a \\ y_{..} &= \sum_{i=1}^{a}\sum_{j=1}^{n} y_{ij} \qquad \bar{y}_{..} = y_{..}/N \end{aligned} \tag{3-3}$$

donde $N = an$ es el número total de observaciones. Se nota que el subíndice "punto" implica la operación suma sobre el subíndice que reemplaza.

El interés se encuentra en probar la igualdad de las a medias de los tratamientos; es decir, $E(y_{ij}) = \mu + \tau_i = \mu_i$, $i = 1, 2, ..., a$. Las hipótesis apropiadas son

$$\begin{aligned} &H_0: \mu_1 = \mu_2 = \ldots = \mu_a \\ &H_1: \mu_i \neq \mu_j \qquad \text{para al menos un par } (i,\ j) \end{aligned}$$

En el modelo de los efectos, la media μ_i del tratamiento i-ésimo se descompone en dos componentes tales que $\mu_i = \mu + \tau_i$. Por lo general, μ se considera como una media global, de tal modo que

$$\frac{\sum_{i=1}^{a} \mu_i}{a} = \mu$$

Esta definición implica que

$$\sum_{i=1}^{a} \tau_i = 0$$

Es decir, los efectos del tratamiento o factor pueden considerarse como desviaciones de la media global.[2] Por consiguiente, una forma equivalente de escribir las hipótesis anteriores es en términos de los efectos de los tratamientos τ_i, por ejemplo

$$H_0: \tau_1 = \tau_2 = \dots \tau_a = 0$$
$$H_1: \tau_i \neq 0 \quad \text{para al menos una } i$$

Por lo tanto, se habla de probar la igualdad de las medias de los tratamientos o de probar que los efectos de los tratamientos (las τ_i) son cero. El procedimiento apropiado para probar la igualdad de las medias de los a tratamientos es el análisis de varianza.

3-3.1 Descomposición de la suma de cuadrados total

El nombre **análisis de varianza** se deriva de la partición de la variabilidad total en sus partes componentes. La suma de cuadrados total corregida

$$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{n} (y_{ij} - \bar{y}_{..})^2$$

se usa como una medida de la variabilidad global de los datos. Intuitivamente, esto es razonable porque, si SS_T tuviera que dividirse por el número apropiado de grados de libertad (en este caso, $an - 1 = N - 1$), se obtendría la **varianza muestral** de las y. La varianza muestral es, desde luego, una medida estándar de variabilidad.

Observe que la suma de cuadrados total corregida SS_T se puede escribir como

$$\sum_{i=1}^{a}\sum_{j=1}^{n} (y_{ij} - \bar{y}_{..})^2 = \sum_{i=1}^{a}\sum_{j=1}^{n} [(\bar{y}_{i.} - \bar{y}_{..}) + (y_{ij} - \bar{y}_{i.})]^2 \tag{3-4}$$

o

$$\begin{aligned}\sum_{i=1}^{a}\sum_{j=1}^{n} (y_{ij} - \bar{y}_{..})^2 = n\sum_{i=1}^{a} (\bar{y}_{i.} - \bar{y}_{..})^2 + \sum_{i=1}^{a}\sum_{j=1}^{n} (y_{ij} - \bar{y}_{i.})^2 \\ + 2\sum_{i=1}^{a}\sum_{j=1}^{n} (\bar{y}_{i.} - \bar{y}_{..})(y_{ij} - \bar{y}_{i.})\end{aligned} \tag{3-5}$$

Sin embargo, el término del producto cruzado de la ecuación 3-5 es cero, ya que

$$\sum_{j=1}^{n} (y_{ij} - \bar{y}_{i.}) = y_{i.} - n\bar{y}_{i.} = y_{i.} - n(y_{i.}/n) = 0$$

Se tiene, por lo tanto,

$$\sum_{i=1}^{a}\sum_{j=1}^{n} (y_{ij} - \bar{y}_{..})^2 = n\sum_{i=1}^{a} (\bar{y}_{i.} - \bar{y}_{..})^2 + \sum_{i=1}^{a}\sum_{j=1}^{n} (y_{ij} - \bar{y}_{i.})^2 \tag{3-6}$$

La ecuación 3-6 establece que puede hacerse la partición de la variabilidad total de los datos, medida por la suma de cuadrados total corregida, en una suma de cuadrados de las diferencias **entre** los promedios de los tratamientos y el gran promedio, más una suma de cuadrados de las diferencias de las observaciones **dentro** de los tratamientos y el promedio de los tratamientos. Entonces, la diferencia entre los promedios

[2] Para más información sobre este tema, referirse al material suplementario del texto del capítulo 3.

de los tratamientos observados y el gran promedio es una medida de las diferencias entre las medias de los tratamientos, mientras que las diferencias de las observaciones dentro de un tratamiento y el promedio del tratamiento, pueden deberse únicamente al error aleatorio. Por lo tanto, la ecuación 3-6 puede escribirse simbólicamente como

$$SS_T = SS_{\text{Tratamientos}} + SS_E$$

donde a $SS_{\text{Tratamientos}}$ se le llama la suma de cuadrados debida a los tratamientos (es decir, entre los tratamientos), y a SS_E se le llama la suma de cuadrados debida al error (es decir, dentro de los tratamientos). Hay $an = N$ observaciones en total; por lo tanto, SS_T tiene $N - 1$ grados de libertad. Hay a niveles del factor (y medias de a tratamientos), de donde $SS_{\text{Tratamientos}}$ tiene $a - 1$ grados de libertad. Por último, dentro de cualquier tratamiento hay n réplicas que proporcionan $n - 1$ grados de libertad con los cuales estimar el error experimental. Puesto que hay a tratamientos, se tienen $a(n - 1) = an - a = N - a$ grados de libertad para el error.

Es útil examinar explícitamente los dos términos del lado derecho de la identidad fundamental del análisis de varianza (ecuación 3-6). Considere la suma de cuadrados del error

$$SS_E = \sum_{i=1}^{a}\sum_{j=1}^{n} (y_{ij} - \bar{y}_{i.})^2 = \sum_{i=1}^{a}\left[\sum_{j=1}^{n} (y_{ij} - \bar{y}_{i.})^2\right]$$

En esta forma es fácil ver que el término entre corchetes, si se divide por $n - 1$, es la varianza muestral del tratamiento i-ésimo, o

$$S_i^2 = \frac{\sum_{j=1}^{n} (y_{ij} - \bar{y}_{i.})^2}{n-1} \qquad i = 1, 2, \ldots, a$$

Ahora pueden combinarse a varianzas muestrales para obtener una sola estimación de la varianza poblacional común de la siguiente manera:

$$\frac{(n-1)S_1^2 + (n-1)S_2^2 + \ldots + (n-1)S_a^2}{(n-1)+(n-1) + \ldots + (n-1)} = \frac{\sum_{i=1}^{a}\left[\sum_{j=1}^{n} (y_{ij} - \bar{y}_{i.})^2\right]}{\sum_{i=1}^{a} (n-1)}$$

$$= \frac{SS_E}{(N-a)}$$

Por lo tanto, $SS_E/(N - a)$ es una estimación combinada de la varianza común dentro de cada uno de los a tratamientos.

De manera similar, si no hubiera diferencias entre las medias de los a tratamientos, podría usarse la variación de los promedios de los tratamientos y el gran promedio para estimar σ^2. Específicamente

$$\frac{SS_{\text{Tratamientos}}}{a-1} = \frac{n\sum_{i=1}^{a} (\bar{y}_{i.} - \bar{y}_{..})^2}{a-1}$$

es una estimación de σ^2 si las medias de los tratamientos son iguales. La razón de esto puede verse de manera intuitiva de la siguiente manera. La cantidad $\Sigma_{i=1}^{a}(\bar{y}_{i.} - \bar{y}_{..})^2/(a - 1)$ estima σ^2/n, la varianza de los pro-

medios de los tratamientos, de donde $n\Sigma_{i=1}^{a}(\bar{y}_{i.} - \bar{y}_{..})^2/(a-1)$ debe estimar σ^2 si no hay diferencias en las medias de los tratamientos.

Se observa que la identidad del análisis de varianza (ecuación 3-6) nos proporciona dos estimaciones de σ^2: una basada en la variabilidad inherente dentro de los tratamientos y una basada en la variabilidad entre los tratamientos. Si no hay diferencias en las medias de los tratamientos, estas dos estimaciones deberán ser muy similares, y si no lo son, se sospecha que la diferencia observada puede ser causada por diferencias en las medias de los tratamientos. Aun cuando se ha usado un razonamiento intuitivo para desarrollar este resultado, puede adoptarse un enfoque un tanto más formal.

A las cantidades

$$MS_{\text{Tratamientos}} = \frac{SS_{\text{Tratamientos}}}{a-1}$$

y

$$MS_E = \frac{SS_E}{N-a}$$

se les llama **cuadrados medios.** Se examinarán ahora los **valores esperados** de estos cuadrados medios. Considere

$$\begin{aligned} E(MS_E) = E\left(\frac{SS_E}{N-a}\right) &= \frac{1}{N-a} E\left[\sum_{i=1}^{a}\sum_{j=1}^{n} (y_{ij} - \bar{y}_{i.})^2\right] \\ &= \frac{1}{N-a} E\left[\sum_{i=1}^{a}\sum_{j=1}^{n} (y_{ij}^2 - 2y_{ij}\bar{y}_{i.} + \bar{y}_{i.}^2)\right] \\ &= \frac{1}{N-a} E\left[\sum_{i=1}^{a}\sum_{j=1}^{n} y_{ij}^2 - 2n\sum_{i=1}^{a} \bar{y}_{i.}^2 + n\sum_{i=1}^{a} \bar{y}_{i.}^2\right] \\ &= \frac{1}{N-a} E\left[\sum_{i=1}^{a}\sum_{j=1}^{n} y_{ij}^2 - \frac{1}{n}\sum_{i=1}^{a} y_{i.}^2\right] \end{aligned}$$

Al sustituir el modelo (ecuación 3-1) en esta ecuación se obtiene

$$E(MS_E) = \frac{1}{N-a} E\left[\sum_{i=1}^{a}\sum_{j=1}^{n} (\mu + \tau_i + \varepsilon_{ij})^2 - \frac{1}{n}\sum_{i=1}^{a}\left(\sum_{j=1}^{n} \mu + \tau_i + \varepsilon_{ij}\right)^2\right]$$

Entonces, al elevar al cuadrado y tomar el valor esperado de la cantidad entre corchetes, se observa que los términos que incluyen a ε_{ij}^2 y $\varepsilon_{i.}^2$ son reemplazados por σ^2 y $n\sigma^2$, respectivamente, debido a que $E(\varepsilon_{ij}) = 0$. Además, todos los productos cruzados que incluyen a ε_{ij} tienen valor esperado cero. Por lo tanto, después de elevar al cuadrado y tomar el valor esperado, la última ecuación se convierte en

$$E(MS_E) = \frac{1}{N-a}\left[N\mu^2 + n\sum_{i=1}^{a} \tau_i^2 + N\sigma^2 - N\mu^2 - n\sum_{i=1}^{a} \tau_i^2 - a\sigma^2\right]$$

o

$$E(MS_E) = \sigma^2$$

Aplicando un enfoque similar puede demostrarse también que[3]

$$E(MS_{\text{Tratamientos}}) = \sigma^2 + \frac{n\sum_{i=1}^{a} \tau_i^2}{a-1}$$

Por lo tanto, como se argumentó heurísticamente, $MS_E = SS_E/(N-a)$ estima σ^2 y, si no hay diferencias en las medias de los tratamientos (lo cual implica que $\tau_i = 0$), $MS_{\text{Tratamientos}} = SS_{\text{Tratamientos}}/(a-1)$ también estima σ^2. Sin embargo, observe que si las medias de los tratamientos difieren, el valor esperado del cuadrado medio de los tratamientos es mayor que σ^2.

Parece claro que es posible realizar una prueba de la hipótesis de que no hay diferencias en las medias de los tratamientos comparando $MS_{\text{Tratamientos}}$ y MS_E. Se considera ahora cómo puede hacerse esta comparación.

3-3.2 Análisis estadístico

Se investiga ahora cómo puede llevarse a cabo una prueba formal de la hipótesis de que no hay diferencias en las medias de los tratamientos (H_0:$\mu_1 = \mu_2 = \cdots = \mu_a$, o de manera equivalente, H_0:$\tau_1 = \tau_2 = \cdots \tau_a = 0$). Puesto que se ha supuesto que los errores ε_{ij} siguen una distribución normal e independiente con media cero y varianza σ^2, las observaciones y_{ij} tienen una distribución normal e independiente con media $\mu + \tau_i$ y varianza σ^2. Por lo tanto, SS_T es una suma de cuadrados de variables aleatorias con una distribución normal; por consiguiente, puede demostrarse que SS_T/σ^2 tiene una distribución ji-cuadrada con $N-1$ grados de libertad. Además, puede demostrarse que SS_E/σ^2 es una variable ji-cuadrada con $N-a$ grados de libertad y que $SS_{\text{Tratamientos}}/\sigma^2$ es una variable ji-cuadrada con $a-1$ grados de libertad si la hipótesis nula H_0:$\tau_i = 0$ es verdadera. Sin embargo, las tres sumas de cuadrados no son necesariamente independientes, ya que la suma de $SS_{\text{Tratamientos}}$ y SS_E es SS_T. El siguiente teorema, que es una forma especial de un teorema atribuido a William Cochran, es útil para establecer la independencia de SS_E y $SS_{\text{Tratamientos}}$.

TEOREMA 3-1 ··

Teorema de Cochran

Sea Z_i igual a NID(0, 1) para $i = 1, 2, ..., \nu$ y

$$\sum_{i=1}^{\nu} Z_i^2 = Q_1 + Q_2 + \ldots + Q_s$$

donde $s \leq \nu$, y Q_i tiene ν_i grados de libertad ($i = 1, 2, ..., s$). Entonces $Q_1, Q_2, ..., Q_s$ son variables aleatorias ji-cuadrada independientes con $\nu_1, \nu_2, ..., \nu_s$ grados de libertad, respectivamente, si y sólo si

$$\nu = \nu_1 + \nu_2 + \ldots + \nu_s$$

··

Puesto que los grados de libertad de $SS_{\text{Tratamientos}}$ y SS_E suman $N-1$, el número total de grados de libertad, el teorema de Cochran implica que $SS_{\text{Tratamientos}}/\sigma^2$ y SS_E/σ^2 son variables aleatorias ji-cuadrada con

[3] Referirse al material suplementario del texto del capítulo 3.

Tabla 3-3 Tabla de análisis de varianza para el modelo con un solo factor y efectos fijos

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0
Entre los tratamientos	$SS_{\text{Tratamientos}} = n\sum_{i=1}^{a} (\bar{y}_{i.} - \bar{y}_{..})^2$	$a-1$	$MS_{\text{Tratamientos}}$	$F_0 = \frac{MS_{\text{Tratamientos}}}{MS_E}$
Error (dentro de los tratamientos)	$SS_E = SS_T - SS_{\text{Tratamientos}}$	$N-a$	MS_E	
Total	$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{n} (y_{ij} - \bar{y}_{..})^2$	$N-1$		

una distribución independiente. Por lo tanto, si la hipótesis nula de que no hay diferencias en las medias de los tratamientos es verdadera, el cociente

$$F_0 = \frac{SS_{\text{Tratamientos}}/(a-1)}{SS_E/(N-a)} = \frac{MS_{\text{Tratamientos}}}{MS_E} \tag{3-7}$$

se distribuye como F con $a-1$ y $N-a$ grados de libertad. La ecuación 3-7 es el **estadístico de prueba** para la hipótesis de que no hay diferencias en las medias de los tratamientos.

Por los cuadrados medios esperados se observa que, en general, MS_E es un estimador insesgado de σ^2. Asimismo, bajo la hipótesis nula, $MS_{\text{Tratamientos}}$ es un estimador insesgado de σ^2. Sin embargo, si la hipótesis nula es falsa, el valor esperado de $MS_{\text{Tratamientos}}$ es mayor que σ^2. Por lo tanto, bajo la hipótesis alternativa, el valor esperado del numerador del estadístico de prueba (ecuación 3-7) es mayor que el valor esperado del denominador, y H_0 deberá rechazarse para valores del estadístico de prueba que son muy grandes. Esto implica una región crítica de una sola cola superior. Por lo tanto, H_0 deberá rechazarse y concluirse que hay diferencias en las medias de los tratamientos si

$$F_0 > F_{\alpha,a-1,N-a}$$

donde F_0 se calcula con la ecuación 3-7. De manera alternativa, podría usarse el enfoque del valor P para tomar una decisión.

Es posible obtener fórmulas para calcular estas sumas de cuadrados reescribiendo y simplificando las definiciones de $SS_{\text{Tratamientos}}$ y SS_T en la ecuación 3-6. Se obtiene así

$$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{n} y_{ij}^2 - \frac{y_{..}^2}{N} \tag{3-8}$$

y

$$SS_{\text{Tratamientos}} = \frac{1}{n}\sum_{i=1}^{a} y_{i.}^2 - \frac{y_{..}^2}{N} \tag{3-9}$$

La suma de cuadrados del error se obtiene por sustracción como

$$SS_E = SS_T - SS_{\text{Tratamientos}} \tag{3-10}$$

El procedimiento de prueba se resume en la tabla 3-3. Se le conoce como **tabla del análisis de varianza.**

EJEMPLO 3-1 ··

El experimento de la resistencia a la tensión

Para ilustrar el análisis de varianza, se retoma al ejemplo que empezó a comentarse en la sección 3-1. Recuerde que al ingeniero de desarrollo de productos le interesa determinar si el peso porcentual del algo-

dón en una fibra sintética afecta la resistencia a la tensión, y ha llevado a cabo un experimento completamente aleatorizado con cinco niveles del peso porcentual del algodón y cinco réplicas. Por conveniencia, a continuación se repiten los datos de la tabla 3-1:

Peso porcentual del algodón	Resistencia a la tensión observada (lb/pulg²)					Totales	Promedios
	1	2	3	4	5	$y_{i.}$	$\bar{y}_{i.}$
15	7	7	15	11	9	49	9.8
20	12	17	12	18	18	77	15.4
25	14	18	18	19	19	88	17.6
30	19	25	22	19	23	108	21.6
35	7	10	11	15	11	54	10.8
						$y_{..} = 376$	$\bar{y}_{..} = 15.04$

Se usará el análisis de varianza para probar H_0:$\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$ contra la hipótesis alternativa H_1:algunas medias son diferentes. Las sumas de cuadrados requeridas se calculan como sigue:

$$SS_T = \sum_{i=1}^{5}\sum_{j=1}^{5} y_{ij}^2 - \frac{y_{..}^2}{N}$$

$$= (7)^2 + (7)^2 + (15)^2 + \cdots + (15)^2 + (11)^2 - \frac{(376)^2}{25} = 636.96$$

$$SS_{\text{Tratamientos}} = \frac{1}{n}\sum_{i=1} y_{i.}^2 - \frac{y_{..}^2}{N}$$

$$= \frac{1}{5}[(49)^2 + \cdots + (54)^2] - \frac{(376)^2}{25} = 475.76$$

$$SS_E = SS_T - SS_{\text{Tratamientos}}$$

$$= 636.96 - 475.76 = 161.20$$

Generalmente estos cálculos se realizarían en una computadora, utilizando paquetes de software con la capacidad de analizar datos de experimentos diseñados.

En la tabla 3-4 se resume el análisis de varianza. Observe que el cuadrado medio entre los tratamientos (118.94) es varias veces mayor que el cuadrado medio dentro de los tratamientos o cuadrado medio del error (8.06). Esto indica que no es posible que las medias de los tratamientos sean iguales. En términos más formales, puede calcularse el cociente F, $F_0 = 118.94/8.06 = 14.76$, y comparar este valor con un punto porcentual apropiado de la cola superior de la distribución $F_{4,20}$. Suponga que el experimentador ha seleccionado $\alpha = 0.05$. En la tabla IV del apéndice se encuentra que $F_{0.05,4,20} = 2.87$. Puesto que $F_0 = 14.76 > 2.87$, se rechaza H_0 y se concluye que las medias de los tratamientos difieren; es decir, el peso porcentual del algodón en la fibra afecta de manera significativa la resistencia a la tensión media. También

Tabla 3-4 Análisis de varianza de los datos de la resistencia a la tensión

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0	Valor P
Peso porcentual del algodón	475.76	4	118.94	F_0=14.76	<0.01
Error	161.20	20	8.06		
Total	636.96	24			

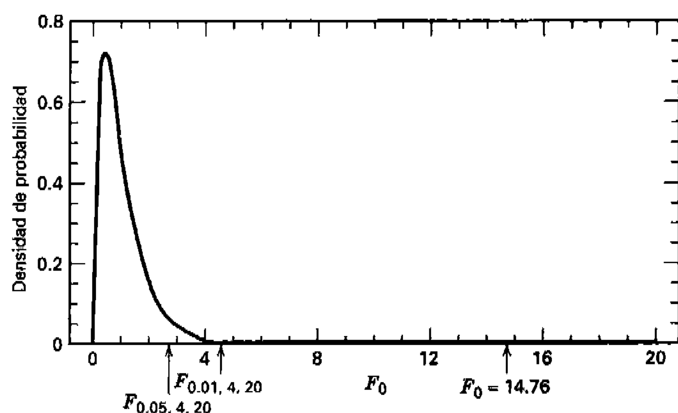


Figura 3-3 La distribución de referencia ($F_{4,\,20}$) para el estadístico de prueba F_0 en el ejemplo 3-1.

podría calcularse un valor P para este estadístico de prueba. En la figura 3-3 se muestra la distribución de referencia ($F_{4,20}$) para el estadístico de prueba F_0. Evidentemente, el valor P es muy pequeño en este caso. Puesto que $F_{0.01,4,20} = 4.43$ y $F_0 > 4.43$, puede concluirse que un límite superior del valor P es 0.01; es decir, $P < 0.01$ (el valor P exacto es $P = 9.11 \times 10^{-6}$).

Cálculos manuales

Posiblemente el lector haya notado que la suma de cuadrados se definió en términos de promedios; es decir, por la ecuación 3-6,

$$SS_{\text{Tratamientos}} = n\sum_{i=1}^{a} (\bar{y}_{i.} - \bar{y}_{..})^2$$

pero las fórmulas de cálculo se desarrollaron utilizando los totales. Por ejemplo, para calcular $SS_{\text{Tratamientos}}$, se usaría la ecuación 3-9:

$$SS_{\text{Tratamientos}} = \frac{1}{n}\sum_{i=1}^{a} y_{i.}^2 - \frac{y_{..}^2}{N}$$

La razón principal de esto es por conveniencia; además, los totales $y_{i.}$ y $y_{..}$ están menos sujetos al error de redondeo que los promedios $\bar{y}_{i.}$ y $\bar{y}_{..}$.

En general, no deberá prestarse demasiada atención a los cálculos, ya que se cuenta con una amplia variedad de programas de computadora para realizarlos. Estos programas de computadora son también útiles para realizar muchos otros análisis asociados con el diseño experimental (como el análisis residual y la verificación de la adecuación del modelo). En muchos casos, estos programas también ayudarán al experimentador a establecer el diseño.

Cuando es necesario realizar los cálculos manualmente, en ocasiones es útil codificar las observaciones. Esto se ilustra en el ejemplo siguiente.

EJEMPLO 3-2

Codificación de observaciones

Los cálculos del análisis de varianza pueden hacerse con frecuencia de manera más precisa o simplificada codificando las observaciones. Por ejemplo, considere los datos de la resistencia a la tensión del ejemplo

Tabla 3-5 Datos codificados de la resistencia a la tensión del ejemplo 3-2

Peso porcentual del algodón	Observaciones					Totales y_i
	1	2	3	4	5	
15	−8	−8	0	−4	−6	−26
20	−3	2	−3	3	3	2
25	−1	3	3	4	4	13
30	4	10	7	4	8	33
35	−8	−5	−4	0	−4	−21

3-1. Suponga que se resta 15 de cada observación. Los datos codificados se muestran en la tabla 3-5. Es sencillo verificar que

$$SS_T = (-8)^2 + (-8)^2 + \cdots + (-4)^2 - \frac{(1)^2}{25} = 636.96$$

$$SS_{\text{Tratamientos}} = \frac{(-26)^2 + (2)^2 + \cdots + (-21)^2}{5} - \frac{(1)^2}{25} = 475.76$$

y

$$SS_E = 161.20$$

Al comparar estas sumas de cuadrados con las que se obtuvieron en el ejemplo 3-1, se observa que al restar una constante de los datos originales las sumas de cuadrados no se modifican.

Suponga ahora que cada una de las observaciones del ejemplo 3-1 se multiplica por 2. Es sencillo verificar que las sumas de cuadrados de los datos transformados son $SS_T = 2547.84$, $SS_{\text{Tratamientos}} = 1903.04$ y $SS_E = 644.80$. Estas sumas de cuadrados parecen diferir considerablemente de las que se obtuvieron en el ejemplo 3-1. Sin embargo, si se dividen por 4 (es decir, 2^2), los resultados son idénticos. Por ejemplo, para la suma de cuadrados de los tratamientos, $1903.04/4 = 475.76$. Asimismo, para los datos codificados, el cociente F es $F = (1903.04/4)/(644.80/20) = 14.76$, que es idéntico al cociente F de los datos originales. Por lo tanto, los análisis de varianza son equivalentes.

Pruebas de aleatorización y análisis de varianza

En el desarrollo del análisis de varianza con la prueba F, se ha utilizado el supuesto de que los errores aleatorios ε_{ij} son variables aleatorias que siguen una distribución normal e independiente. También es posible justificar la prueba F como la aproximación de una **prueba de aleatorización.** Para ilustrar esto, suponga que se tienen cinco observaciones de cada uno de dos tratamientos y que quiere probarse la igualdad de las medias de los tratamientos. Los datos aparecerían así:

Tratamiento 1	Tratamiento 2
y_{11}	y_{21}
y_{12}	y_{22}
y_{13}	y_{23}
y_{14}	y_{24}
y_{15}	y_{25}

Podría usarse el análisis de varianza con la prueba F para probar $H_0{:}\mu_1 = \mu_2$. De manera alternativa, podría recurrirse a un enfoque un tanto diferente. Suponga que se consideran todas las formas posibles de

asignar los 10 números de la muestra anterior a los dos tratamientos. Hay 10!/5!5! = 252 arreglos posibles de las 10 observaciones. Si no hay ninguna diferencia en las medias de los tratamientos, los 252 arreglos son igualmente posibles. Para cada uno de los 252 arreglos, se calcula el valor del estadístico F usando la ecuación 3-7. A la distribución de estos valores F se le llama **distribución de aleatorización,** y un valor grande de F indica que los datos no son consistentes con la hipótesis H_0:$\mu_1 = \mu_2$. Por ejemplo, si el valor de F que se observó realmente fue excedido sólo por 5 de los valores de la distribución de aleatorización, esto correspondería con el rechazo de H_0:$\mu_1 = \mu_2$ con un nivel de significación de $\alpha = 5/252 = 0.0198$ (o 1.98%). Observe que no es necesario ningún supuesto de normalidad en este enfoque.

La dificultad con este enfoque es que, incluso en problemas relativamente pequeños, los cálculos requeridos hacen inviable la enumeración de la distribución de aleatorización exacta. Sin embargo, numerosos estudios han demostrado que la distribución F común de la teoría normal es una buena aproximación de la distribución de aleatorización exacta. Por lo tanto, incluso sin el supuesto de normalidad, la prueba F puede considerarse como una aproximación de la prueba de aleatorización. Para más detalles sobre las pruebas de aleatorización en el análisis de varianza, ver Box, Hunter y Hunter [18].

3-3.3 Estimación de los parámetros del modelo

Se presentan ahora los estimadores de los parámetros del modelo con un solo factor

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij}$$

y los intervalos de confianza para las medias de los tratamientos. Más adelante se demostrará que estimadores razonables de la media global y de los efectos de los tratamientos están dados por

$$\begin{aligned} \hat{\mu} &= \bar{y}_{..} \\ \hat{\tau}_i &= \bar{y}_{i.} - \bar{y}_{..}, \qquad i = 1, 2, \ldots, a \end{aligned} \tag{3-11}$$

Estos estimadores poseen un considerable atractivo intuitivo; observe que la media global se estima con el gran promedio de las observaciones y que el efecto de cualquier tratamiento no es sino la diferencia entre el promedio del tratamiento y el gran promedio.

Es posible determinar con facilidad una estimación del **intervalo de confianza** de la media del tratamiento i-ésimo. La media del tratamiento i-ésimo es

$$\mu_i = \mu + \tau_i$$

Un estimador puntual de μ_i sería $\hat{\mu}_i = \hat{\mu} + \hat{\tau}_i = \bar{y}_{i.}$. Ahora bien, si se supone que los errores siguen una distribución normal, cada $\bar{y}_{i.}$ es una NID(μ_i, σ^2/n). Por lo tanto, si σ^2 fuera conocida, podría usarse la distribución normal para definir el intervalo de confianza. Al utilizar MS_E como estimador de σ^2, el intervalo de confianza se basaría en la distribución t. Por lo tanto, un intervalo de confianza de $100(1 - \alpha)$ por ciento para la media μ_i del tratamiento i-ésimo es

$$\bar{y}_{i.} - t_{\alpha/2,N-a}\sqrt{\frac{MS_E}{n}} \le \mu_i \le \bar{y}_{i.} + t_{\alpha/2,N-a}\sqrt{\frac{MS_E}{n}} \tag{3-12}$$

Un intervalo de confianza de $100(1 - \alpha)$ por ciento para la diferencia en las medias de dos tratamientos cualesquiera, por ejemplo $\mu_i - \mu_j$, sería

$$\bar{y}_{i.} - \bar{y}_{j.} - t_{\alpha/2,N-a}\sqrt{\frac{2MS_E}{n}} \le \mu_i - \mu_j \le \bar{y}_{i.} - \bar{y}_{j.} + t_{\alpha/2,N-a}\sqrt{\frac{2MS_E}{n}} \tag{3-13}$$

EJEMPLO 3-3 ···

Utilizando los datos del ejemplo 3-1 pueden encontrarse las estimaciones de la media global y de los efectos de los tratamientos como $\hat{\mu} = 376/25 = 15.04$ y

$$\hat{\tau}_1 = \bar{y}_{1.} - \bar{y}_{..} = 9.80 - 15.04 = -5.24$$
$$\hat{\tau}_2 = \bar{y}_{2.} - \bar{y}_{..} = 15.40 - 15.04 = +0.36$$
$$\hat{\tau}_3 = \bar{y}_{3.} - \bar{y}_{..} = 17.60 - 15.04 = -2.56$$
$$\hat{\tau}_4 = \bar{y}_{4.} - \bar{y}_{..} = 21.60 - 15.04 = +6.56$$
$$\hat{\tau}_5 = \bar{y}_{5.} - \bar{y}_{..} = 10.80 - 15.04 = -4.24$$

Un intervalo de confianza de 95% para la media del tratamiento 4 (30% de algodón) se calcula con la ecuación 3-12 como

$$21.60 - 2.086\sqrt{\frac{8.06}{5}} \le \mu_4 \le 21.60 + 2.086\sqrt{\frac{8.06}{5}}$$

o

$$21.60 - 2.65 \le \mu_4 \le 21.60 + 2.65$$

Por lo tanto, el intervalo de confianza de 95% deseado es $18.95 \le \mu_4 \le 24.25$.

···

Intervalos de confianza simultáneos

Las expresiones para los intervalos de confianza dados en las ecuaciones 3-12 y 3-13 son intervalos de confianza **uno a la vez.** Es decir, el nivel de confianza $1 - \alpha$ sólo se aplica a una estimación particular. Sin embargo, en muchos problemas, el experimentador tal vez quiera calcular varios intervalos de confianza, uno para cada una de varias medias o diferencias entre medias. Si hay interés en r de estos intervalos de confianza de $100(1-\alpha)$ por ciento, la probabilidad de que los r intervalos sean correctos **simultáneamente** es al menos $1 - r\alpha$. A la probabilidad $r\alpha$ se le llama con frecuencia **índice de error en el modo del experimento** o coeficiente de confianza global. El número de intervalos r no tiene que ser muy grande antes de que el conjunto de intervalos de confianza se vuelva relativamente falto de información. Por ejemplo, si hay $r = 5$ intervalos y $\alpha = 0.05$ (una elección típica), el nivel de confianza simultáneo para el conjunto de los cinco intervalos de confianza es de al menos 0.75, y si $r = 10$ y $\alpha = 0.05$, el nivel de confianza simultáneo es de al menos 0.50.

Un enfoque para asegurarse de que el nivel de confianza simultáneo no sea demasiado pequeño es sustituir $\alpha/2$ con $\alpha/(2r)$ en las ecuaciones 3-12 y 3-13 del intervalo de confianza uno a la vez. A éste se le llama el **método de Bonferroni,** y le permite al experimentador construir un conjunto de r intervalos de confianza simultáneos para las medias de los tratamientos o las diferencias en las medias de los tratamientos para los que el nivel de confianza global es de al menos $100(1-\alpha)$ por ciento. Cuando r no es muy grande, éste es un método muy atinado que produce intervalos de confianza razonablemente cortos. Para más información, referirse al material suplementario del texto del capítulo 3.

3-3.4 Datos no balanceados

En algunos experimentos con un solo factor, puede ser diferente el número de observaciones que se hacen dentro de cada tratamiento. Se dice entonces que el diseño es **no balanceado.** Sigue siendo posible aplicar el análisis de varianza descrito arriba, pero deben hacerse ligeras modificaciones en las fórmulas

de las sumas de cuadrados. Sea que se hagan n_i observaciones bajo el tratamiento i ($i = 1, 2, ..., a$) y que $N = \Sigma_{i=1}^{a} n_i$. Las fórmulas para calcular manualmente SS_T y $SS_{\text{Tratamientos}}$ quedan como

$$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{n_i} y_{ij}^2 - \frac{y_{..}^2}{N} \tag{3-14}$$

y

$$SS_{\text{Tratamientos}} = \sum_{i=1}^{a} \frac{y_{i.}^2}{n_i} - \frac{y_{..}^2}{N} \tag{3-15}$$

No se requieren más cambios en el análisis de varianza.

Hay dos ventajas al elegir un diseño balanceado. Primera, el estadístico de prueba es relativamente insensible a las desviaciones pequeñas del supuesto de la igualdad de las varianzas de los a tratamientos cuando los tamaños de las muestras son iguales. No es éste el caso cuando los tamaños de las muestras son diferentes. Segunda, la potencia de la prueba se maximiza cuando las muestras tienen el mismo tamaño.

3-4 VERIFICACIÓN DE LA ADECUACIÓN DEL MODELO

La descomposición de la variabilidad presente en las observaciones mediante la identidad del análisis de varianza (ecuación 3-6) es una relación puramente algebraica. Sin embargo, el uso de la partición para probar formalmente que no hay diferencias en las medias de los tratamientos requiere que se satisfagan ciertos supuestos. Específicamente, estos supuestos son que el modelo

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij}$$

describe de manera adecuada las observaciones, y que los errores siguen una distribución normal e independiente con media cero y varianza σ^2 constante pero desconocida. Si estos supuestos se satisfacen, el procedimiento del análisis de varianza es una prueba exacta de la hipótesis de que no hay diferencias en las medias de los tratamientos.

Sin embargo, es común que en la práctica estos supuestos no se satisfagan exactamente. Por consiguiente, en general no es prudente confiar en el análisis de varianza hasta haber verificado estos supuestos. Las violaciones de los supuestos básicos y la adecuación del modelo pueden investigarse con facilidad mediante el examen de los **residuales.** El residual de la observación j-ésima en el tratamiento i-ésimo se define como

$$e_{ij} = y_{ij} - \hat{y}_{ij} \tag{3-16}$$

donde $\hat{y}_{ij}$ es una estimación de la observación y_{ij} correspondiente que se obtiene como sigue:

$$\begin{aligned}\hat{y}_{ij} &= \hat{\mu} + \hat{\tau}_i \\ &= \bar{y}_{..} + (\bar{y}_{i.} - \bar{y}_{..}) \\ &= \bar{y}_{i.}\end{aligned} \tag{3-17}$$

La ecuación 3-17 da el resultado intuitivamente claro de que la estimación de cualquier observación en el tratamiento i-ésimo no es sino el promedio del tratamiento correspondiente.

El examen de los residuales deberá ser una parte automática de cualquier análisis de varianza. Si el modelo es adecuado, los residuales deberán estar **sin estructura;** es decir, no deberán contener patrones obvios. A través de un estudio de los residuales, pueden descubrirse muchos tipos de inadecuaciones del modelo y violaciones de los supuestos subyacentes. En esta sección se indica cómo puede hacerse con facilidad la verificación de diagnóstico del modelo mediante el análisis gráfico de los residuales y cómo resolver varias anormalidades que ocurren comúnmente.

3-4.1 El supuesto de normalidad

La verificación del supuesto de normalidad podría hacerse graficando un histograma de los residuales. Si se satisface el supuesto de NID$(0, \sigma^2)$ para los errores, esta gráfica deberá aparecer como una muestra de una distribución normal con centro en cero. Desafortunadamente, cuando se trabaja con muestras pequeñas, suelen ocurrir fluctuaciones significativas, por lo que la aparición de una desviación moderada de la normalidad no implica necesariamente una violación seria de los supuestos. Las desviaciones marcadas de la normalidad son potencialmente serias y requieren análisis adicional.

Un procedimiento en extremo útil es construir una gráfica de probabilidad normal de los residuales. Recuerde que en el capítulo 2 se utilizó una gráfica de probabilidad normal de los datos originales para verificar el supuesto de normalidad cuando se usó la prueba t. En el análisis de varianza, por lo general es más eficaz (y directo) hacer lo mismo con los residuales. Si la distribución fundamental de los errores es normal, esta gráfica tendrá la apariencia de una línea recta. Para visualizar la línea recta, deberá prestarse más atención a los valores centrales de la gráfica que a los valores extremos.

En la tabla 3-6 se muestran los datos originales y los residuales de los datos de la resistencia a la tensión del ejemplo 3-1. La gráfica de probabilidad normal se muestra en la figura 3-4. La impresión general que surge al examinar esta representación es que la distribución de los errores puede tener un ligero sesgo, con la cola derecha siendo más larga que la izquierda. La tendencia de la gráfica de probabilidad normal a curvarse hacia abajo ligeramente del lado izquierdo, implica que la cola izquierda de la distribución de los errores sea un tanto más *delgada* de lo que se anticiparía con una distribución normal; es decir, los residuales negativos no son tan grandes (en valor absoluto) como se esperaba. Sin embargo, esta gráfica no muestra una desviación marcada de la distribución normal.

En general, las desviaciones moderadas de la normalidad no son motivo de gran preocupación en el análisis de varianza de efectos fijos (recuerde el análisis de las pruebas de aleatorización de la sección 3-3.2). Una distribución de los errores que tiene colas considerablemente más gruesas o delgadas que la distribución normal es motivo de mayor preocupación que una distribución sesgada. Puesto que la prueba F sólo se afecta ligeramente, se dice que el análisis de varianza (y los procedimientos relacionados como las comparaciones múltiples) es **robusto** con respecto al supuesto de normalidad. Las desviaciones de la normalidad hacen por lo general que tanto el verdadero nivel de significación como la verdadera potencia difieran ligeramente de los valores anunciados, con la potencia siendo generalmente más baja. El modelo de los efectos aleatorios que se revisará en el capítulo 12 se afecta en forma más severa por la no normalidad.

Tabla 3-6 Datos y residuales del ejemplo 3-1[a]

Peso porcentual del algodón	Observaciones (j)					
	1	2	3	4	5	$\hat{y}_{ij} = \bar{y}_{i}$
15	−2.8 7 (15)	−2.8 7 (19)	5.2 15 (25)	1.2 11 (12)	−0.8 9 (6)	9.8
20	−3.4 12 (8)	1.6 17 (14)	−3.4 12 (1)	2.6 18 (11)	2.6 19 (3)	15.4
25	−3.6 14 (18)	0.4 18 (13)	0.4 18 (20)	1.4 19 (7)	1.4 19 (9)	17.6
30	−2.6 19 (22)	3.4 25 (5)	0.4 22 (2)	−2.6 19 (24)	1.4 23 (10)	21.6
35	−3.8 7 (17)	−0.8 10 (21)	0.2 11 (4)	4.2 15 (16)	0.2 11 (23)	10.8

[a]Los residuales se indican en el recuadro de cada celda. Los números entre paréntesis indican el orden en que se recolectaron los datos.

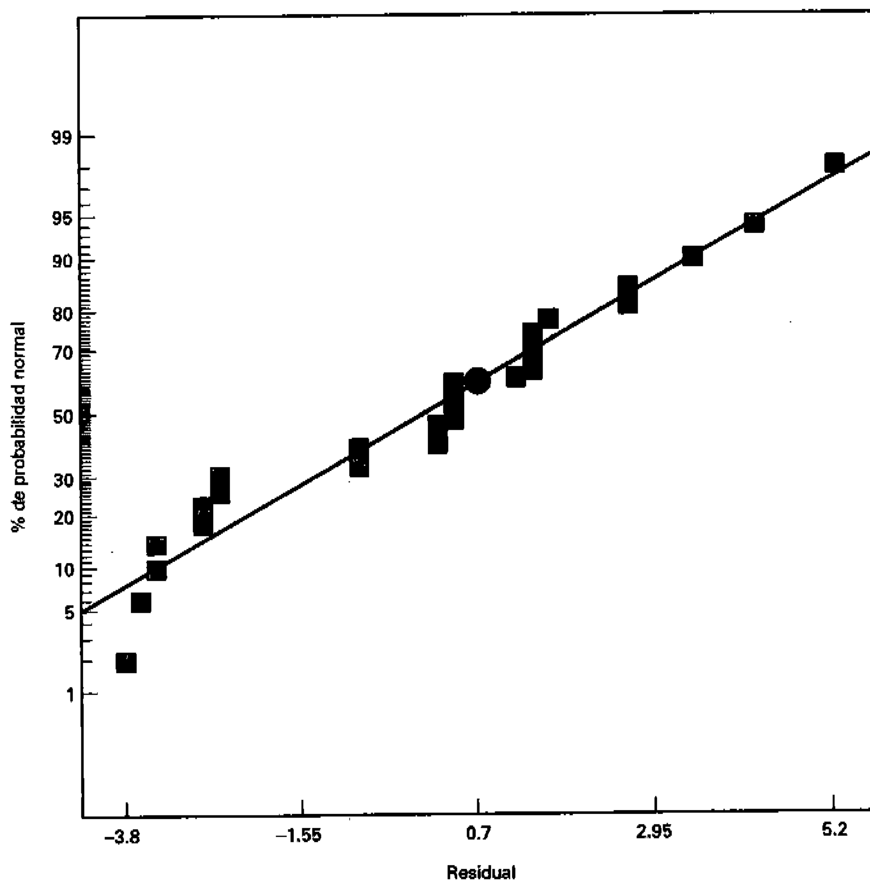


Figura 3-4 Gráfica de probabilidad normal de los residuales del ejemplo 3-1.

Una anomalía muy común que suele ponerse de manifiesto en las gráficas de probabilidad normal es un residual que es mucho más grande que cualquier otro. A un residual así se le llama con frecuencia **punto atípico.** La presencia de uno o más puntos atípicos puede introducir serias distorsiones en el análisis de varianza, por lo que cuando se localiza un punto atípico potencial, se requiere una investigación atenta. En muchas ocasiones, la causa del punto atípico es un error en los cálculos o un error al codificar o copiar los datos. Si no es ésta la causa, las circunstancias experimentales que rodean esta corrida particular deben estudiarse con atención. Si la respuesta atípica ocurre en un valor particularmente deseable (alta resistencia, costo bajo, etc.), el punto atípico puede ser más informativo que el resto de los datos. Deberá tenerse cuidado de no rechazar o descartar una observación atípica a menos que se tengan razones no estadísticas de peso para hacerlo. En el peor de los casos, puede terminarse con dos análisis; uno con el punto atípico y uno sin él.

Existen varios procedimientos estadísticos formales para detectar puntos atípicos (por ejemplo, ver Barnett y Lewis [8], John y Prescott [60] y Stefansky [107]). Puede hacerse una verificación aproximada de los puntos atípicos examinando los **residuales estandarizados**

$$d_{ij} = \frac{e_{ij}}{\sqrt{MS_E}} \tag{3-18}$$

Si los errores ε_{ij} son $N(0, \sigma^2)$, los residuales estandarizados deberán ser aproximadamente normales con media cero y varianza unitaria. Por lo tanto, cerca de 68% de los residuales estandarizados deberán estar incluidos dentro de los límites ±1, cerca de 95% de ellos deberán estar incluidos dentro de ±2 y virtualmente todos ellos deberán estar incluidos dentro de ±3. Un residual mayor que 3 o 4 desviaciones estándar a partir de cero es un punto atípico potencial.

Para los datos de la resistencia a la tensión del ejemplo 3-1, la gráfica de probabilidad normal no produce indicio alguno de puntos atípicos. Además, el residual estandarizado mayor es

$$d_{13} = \frac{e_{13}}{\sqrt{MS_E}} = \frac{5.2}{\sqrt{8.06}} = \frac{5.2}{2.84} = 1.83$$

el cual no deberá ser motivo de preocupación.

3-4.2 Gráfica de los residuales en secuencia en el tiempo

La graficación de los residuales en el orden temporal de la recolección de los datos es útil para detectar **correlaciones** entre los residuales. Una tendencia a tener corridas de residuales positivos y negativos indica una correlación positiva. Esto implicaría que el **supuesto de independencia** de los errores ha sido violado. Se trata de un problema potencialmente serio y cuya solución es difícil, por lo que de ser posible es importante evitar el problema cuando se colecten los datos. La aleatorización adecuada del experimento es un paso importante para conseguir la independencia.

En ocasiones las habilidades del experimentador (o los sujetos) pueden cambiar conforme el experimento avanza, o el proceso bajo estudio puede "vagar sin rumbo" o volverse más errático. Esto producirá con frecuencia un cambio en la varianza del error con el tiempo. Esta condición suele llevar a una gráfica de los residuales contra el tiempo que muestra una dispersión mayor en uno de sus extremos que en el otro. Una varianza no constante es un problema potencialmente serio. En las secciones 3-4.3 y 3-4.4 se abundará sobre el tema.

En la tabla 3-6 se muestran los residuales y la secuencia en el tiempo de la recolección de los datos para el experimento de la resistencia a la tensión. En la figura 3-5 se presenta una gráfica de estos residua-

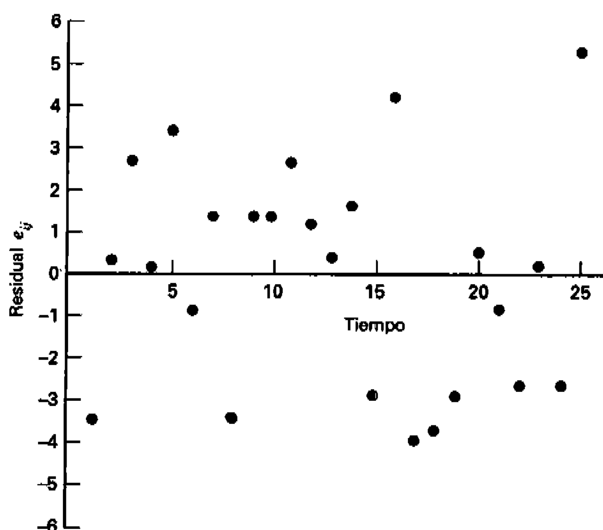


Figura 3-5 Gráfica de los residuales contra el tiempo.

les contra el tiempo. No hay razón para sospechar cualquier violación de los supuestos de independencia o de una varianza constante.

3-4.3 Gráfica de los residuales contra los valores ajustados

Si el modelo es correcto y se satisfacen los supuestos, los residuales deberán estar sin estructura; en particular, no deberán estar relacionados con ninguna otra variable, incluyendo la respuesta predicha. Una verificación simple es graficar los residuales contra los valores ajustados $\hat{y}_{ij}$. (Para el modelo de un experimento con un solo factor, recuerde que $\hat{y}_{ij} = \bar{y}_{i.}$, el promedio del tratamiento i-ésimo.) Esta gráfica no deberá mostrar ningún patrón obvio. En la figura 3-6 se grafican los residuales contra los valores ajustados para los datos de la resistencia a la tensión del ejemplo 3-1. No es evidente ninguna estructura inusual.

Un defecto que sale a relucir en ocasiones en esta gráfica es una **varianza no constante.** En ocasiones la varianza de las observaciones se incrementa cuando la magnitud de la observación se incrementa. Éste sería el caso si el error o ruido de fondo del experimento fuera un porcentaje constante de la magnitud de la observación. (Esto ocurre comúnmente con muchos instrumentos de medición; el error es un porcentaje de la escala de medición.) Si éste fuera el caso, los residuales se harían mayores conforme y_{ij} se hiciera más grande, y la gráfica de los residuales contra y_{ij} se vería como un embudo o un megáfono con la boca hacia afuera. Una varianza no constante también surge en los casos en que los datos siguen una distribución no normal, sesgada, porque en las distribuciones sesgadas la varianza tiende a ser una función de la media.

Si se viola el supuesto de la homogeneidad de las varianzas, la prueba F sólo resulta afectada ligeramente en el modelo balanceado (el mismo tamaño de la muestra en todos los tratamientos) con efectos fijos. Sin embargo, en diseños no balanceados o en casos en que una de las varianzas es considerablemente más grande que las demás, el problema es más grave. Específicamente, si los niveles del factor que tienen las varianzas mayores corresponden también con los tamaños de las muestras más pequeños, el índice de error tipo I real es mayor que lo previsto (o los intervalos de confianza tienen niveles de confianza reales más bajos que los que fueron especificados). Recíprocamente, si los niveles del factor con las varianzas mayores tienen también los tamaños de las muestras mayores, los niveles de significación son mucho menores que lo anticipado (los niveles de confianza son más altos). Ésta es una buena razón para escoger **tamaños de las muestras iguales** siempre que sea posible. Para los modelos con efectos aleatorios, las

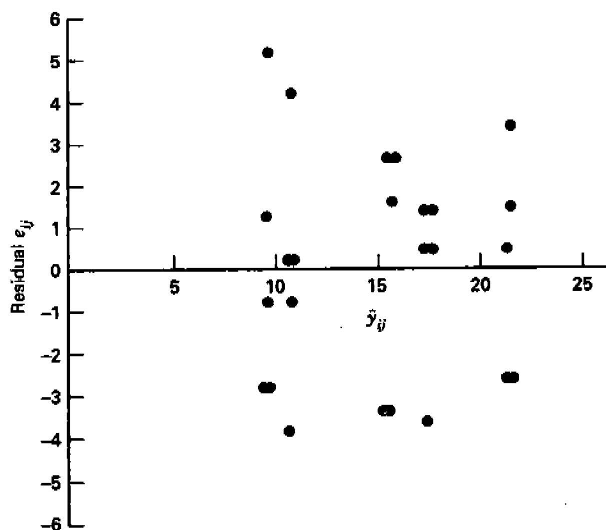


Figura 3-6 Gráfica de los residuales contra los valores ajustados.

varianzas del error diferentes pueden introducir alteraciones significativas en las inferencias sobre los componentes de la varianza, incluso cuando se usan diseños balanceados.

El enfoque usual para abordar el problema de una varianza no constante que ocurre por las razones expuestas antes consiste en aplicar una **transformación para estabilizar la varianza** para correr después el análisis de varianza en los datos transformados. En este enfoque, deberá tenerse presente que las conclusiones del análisis de varianza se aplican a las poblaciones *transformadas.*

Se han dedicado considerables esfuerzos de investigación a la selección de una transformación adecuada. Si los experimentadores conocen la distribución teórica de las observaciones, pueden hacer uso de esta información para elegir la transformación. Por ejemplo, si las observaciones siguen la distribución de Poisson, se usaría la **transformación de la raíz cuadrada** $y_{ij}^* = \sqrt{y_{ij}}$ o $y_{ij}^* = \sqrt{1+y_{ij}}$. Si los datos siguen la distribución lognormal, la **transformación logarítmica** $y_{ij}^* = \log y_{ij}$ es adecuada. Para datos binomiales expresados como fracciones, la **transformación arcsen** $y_{ij}^* = \text{arcsen}\sqrt{y_{ij}}$ es útil. Cuando no hay una transformación obvia, el experimentador realizará por lo general la búsqueda *empírica* de una transformación que iguale la varianza independientemente del valor de la media. Al final de esta sección se ofrecen algunos consejos al respecto. En los experimentos factoriales, los cuales se introducen en el capítulo 5, otro enfoque es seleccionar una transformación que minimice el cuadrado medio de las interacciones, siendo el resultado un experimento cuya interpretación es más sencilla. En el capítulo 14 se revisan con mayor detalle los métodos para seleccionar analíticamente la forma de la transformación. Las transformaciones que se hacen para la desigualdad de la varianza afectan también la forma de la distribución del error. En la mayoría de los casos, la transformación hace que la distribución del error esté más cerca de la distribución normal. Para más detalles sobre las transformaciones, referirse a Bartlett [7], Box y Cox [15], Dolby [38] y Draper y Hunter [39].

Pruebas estadísticas para la igualdad de la varianza

Aun cuando es frecuente el uso de las gráficas residuales para diagnosticar la desigualdad de la varianza, se han propuesto también varias pruebas estadísticas. Estas pruebas pueden considerarse como pruebas formales de las hipótesis

$$H_0: \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_a^2$$

$$H_1: \text{el enunciado anterior no es verdadero para al menos una } \sigma_i^2$$

Un procedimiento muy utilizado es la **prueba de Bartlett.** El procedimiento incluye el cálculo de un estadístico cuya distribución de muestreo está aproximada muy de cerca por la distribución ji-cuadrada con $a-1$ grados de libertad cuando las a muestras aleatorias son de poblaciones normales independientes. El estadístico de prueba es

$$\chi_0^2 = 2.3026\frac{q}{c} \tag{3-19}$$

donde

$$q = (N-a)\log_{10} S_p^2 - \sum_{i=1}^{a} (n_i - 1)\log_{10} S_i^2$$

$$c = 1 + \frac{1}{3(a-1)}\left(\sum_{i=1}^{a} (n_i-1)^{-1} - (N-a)^{-1}\right)$$

$$S_p^2 = \frac{\sum_{i=1}^{a} (n_i-1)S_i^2}{N-a}$$

y S_i^2 es la varianza muestral de la población i-ésima.

La cantidad q es grande cuando la diferencia entre las varianzas muestrales S_i^2 es considerablemente grande, y es igual a cero cuando todas las S_i^2 son iguales. Por lo tanto, H_0 deberá rechazarse para los valores de χ_0^2 que sean muy grandes; es decir, se rechaza H_0 sólo cuando

$$\chi_0^2 > \chi_{\alpha,a-1}^2$$

donde $\chi_{\alpha,a-1}^2$ es el punto porcentual α superior de la distribución ji-cuadrada con $a-1$ grados de libertad. También podría usarse el enfoque del valor P para tomar una decisión.

La prueba de Bartlett es muy sensible al supuesto de normalidad. Por consiguiente, cuando la validez de este supuesto está en duda, no deberá usarse la prueba de Bartlett.

EJEMPLO 3-4 ....

Ya que el supuesto de normalidad no está en entredicho, la prueba de Bartlett puede aplicarse a los datos de la resistencia a la tensión del experimento del peso porcentual de algodón del ejemplo 3-1. Se calculan primero las varianzas muestrales de cada tratamiento y se encuentra que $S_1^2 = 11.2, S_2^2 = 9.8, S_3^2 = 4.3, S_4^2 = 6.8$ y $S_5^2 = 8.2$. Entonces

$$S_p^2 = \frac{4(11.2)+4(9.8)+4(4.3)+4(6.8)+4(8.2)}{20} = 8.06$$

$$q = 20\log_{10}(8.06) - 4[\log_{10} 11.2 + \log_{10} 9.8 + \log_{10} 4.3 + \log_{10} 6.8 + \log_{10} 8.2] = 0.45$$

$$c = 1 + \frac{1}{3(4)}\left(\frac{5}{4} - \frac{1}{20}\right) = 1.10$$

y el estadístico de prueba es

$$\chi_0^2 = 2.3026\frac{(0.45)}{(1.10)} = 0.93$$

Puesto que $\chi_{0.05,4}^2 = 9.49$, no puede rechazarse la hipótesis nula y se concluye que las cinco varianzas son iguales. Se trata de la misma conclusión a la que se llegó al analizar la gráfica de los residuales contra los valores ajustados.

....

Debido a que la prueba de Bartlett es sensible al supuesto de normalidad, puede haber situaciones en las que sería útil un procedimiento alternativo. Anderson y McLean [2] presentan una atinada revisión de algunas pruebas estadísticas de la igualdad de la varianza. La **prueba de Levene modificada** (ver Levene [72] y Conover, Johnson y Johnson [31]) es un procedimiento muy útil que es robusto en cuanto a las desviaciones de la normalidad. Para probar la hipótesis de que las varianzas son iguales en todos los tratamientos, la prueba de Levene modificada utiliza la desviación absoluta de las observaciones y_{ij} de cada tratamiento de la mediana de los tratamientos, por ejemplo $\tilde{y}_i$. Sea que estas desviaciones se denoten por

$$d_{ij} = |y_{ij} - \tilde{y}_i| \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, n_i \end{cases}$$

La prueba de Levene modificada evalúa entonces si la media de estas desviaciones es igual o no para todos los tratamientos. Cuando las desviaciones medias son iguales, las varianzas de las observaciones de

Tabla 3-7 Datos de la descarga pico

Método de estimación	Observaciones						$\bar{y}_{i.}$	$\tilde{y}_i$	S_i
1	0.34	0.12	1.23	0.70	1.75	0.12	0.71	0.520	0.66
2	0.91	2.94	2.14	2.36	2.86	4.55	2.63	2.610	1.09
3	6.31	8.37	9.75	6.09	9.82	7.24	7.93	7.805	1.66
4	17.15	11.82	10.95	17.20	14.35	16.82	14.72	15.59	2.77

Método de estimación	Desviaciones d_{ij} para la prueba de Levene modificada					
1	0.18	0.40	0.71	0.18	1.23	0.40
2	1.70	0.33	0.47	0.25	0.25	1.94
3	1.495	0.565	1.945	1.715	2.015	0.565
4	1.56	3.77	4.64	1.61	1.24	1.23

todos los tratamientos serán iguales. El estadístico de prueba para la prueba de Levene es simplemente el estadístico F ANOVA usual para probar la igualdad de las medias que se aplica a las desviaciones absolutas.

EJEMPLO 3-5 ··························

Un ingeniero civil está interesado en determinar si cuatro métodos diferentes para estimar la frecuencia de las inundaciones producen estimaciones equivalentes de la descarga pico cuando se aplican a la misma cuenca. Cada procedimiento se usa seis veces en la cuenca, y los datos de las descargas resultantes (en pies cúbicos por segundo) se muestran en la parte superior de la tabla 3-7. El análisis de varianza de los datos, el cual se resume en la tabla 3-8, implica que hay una diferencia en las estimaciones de la descarga pico promedio obtenidas en los cuatro procedimientos. La gráfica de los residuales contra los valores ajustados, la cual se muestra en la figura 3-7, es preocupante porque la forma de embudo con la boca hacia afuera indica que no se satisface el supuesto de una varianza constante.

Se aplicará la prueba de Levene modificada a los datos de la descarga pico. La parte superior de la tabla 3-7 contiene las medianas de los tratamientos $\tilde{y}_i$ y la parte inferior contiene las desviaciones d_{ij} alrededor de las medianas. La prueba de Levene consiste en realizar un análisis de varianza estándar en las d_{ij}. El estadístico de prueba F que resulta en este caso es $F_0 = 4.55$, para el cual el valor P es $P = 0.0137$. Por lo tanto, la prueba de Levene rechaza la hipótesis nula de que las varianzas son iguales, coincidiendo en esencia con el diagnóstico que se hizo a partir del examen visual de la figura 3-7. Los datos de la descarga pico son un buen candidato para una transformación de datos.

··························

Selección empírica de una transformación

Se señaló ya que si los experimentadores conocieran la relación entre la varianza de las observaciones y la media, podrían usar esta información como guía para la selección de la forma de la transformación. Se

Tabla 3-8 Análisis de varianza de los datos de la descarga pico

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0	Valor P
Métodos	708.3471	3	236.1157	76.07	<0.001
Error	62.0811	20	3.1041		
Total	770.4282	23			

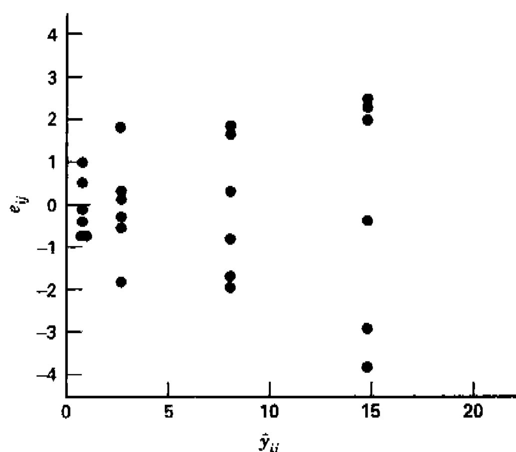


Figura 3-7 Gráfica de los residuales contra $\hat{y}_{ij}$ para el ejemplo 3-5.

desarrolla ahora este punto y se presenta un método para seleccionar empíricamente la forma de la transformación requerida de los datos.

Sea $E(y) = \mu$ la media de y, y suponga que la desviación estándar de y es proporcional a una potencia de la media de y tal que

$$\sigma_y \propto \mu^{\alpha}$$

Quiere encontrarse una transformación de y que produzca una varianza constante. Suponga que la transformación es una potencia de los datos originales, por ejemplo

$$y^* = y^{\lambda} \tag{3-20}$$

Puede demostrarse entonces que

$$\sigma_{y^*} \propto \mu^{\lambda+\alpha-1} \tag{3-21}$$

Evidentemente, si se hace $\lambda = 1 - \alpha$, la varianza de los datos transformados y^* es constante.

En la tabla 3-9 se resumen varias de las transformaciones comunes analizadas anteriormente. Observe que $\lambda = 0$ implica la transformación logarítmica. Estas transformaciones se enlistan en el orden de **fuerza** creciente. Por fuerza de una transformación se entiende la cantidad de curvatura que induce. Una transformación suave aplicada a datos que se extienden en un rango estrecho tiene escaso efecto sobre el

Tabla 3-9 Transformaciones para estabilizar la varianza

Relación entre σ_y y μ	α	$\lambda = 1 - \alpha$	Transformación	Comentario
$\sigma_y \propto$ constante	0	1	Sin transformación	
$\sigma_y \propto \mu^{1/2}$	1/2	1/2	Raíz cuadrada	Datos (números) de Poisson
$\sigma_y \propto \mu$	1	0	Log	
$\sigma_y \propto \mu^{3/2}$	3/2	−1/2	Raíz cuadrada recíproca	
$\sigma_y \propto \mu^2$	2	−1	Recíproco	

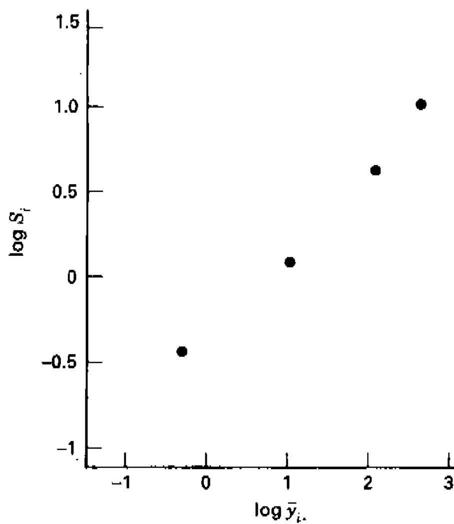


Figura 3-8 Gráfica de log S_i contra log$\bar{y}_{i.}$ para los datos de la descarga pico del ejemplo 3-5.

análisis, mientras que una transformación fuerte aplicada en un rango amplio puede tener resultados dramáticos. Con frecuencia las transformaciones tienen escaso efecto a menos que el cociente $y_{máx}/y_{mín}$ sea mayor que 2 o 3.

En muchas situaciones de diseño experimental en las que se usan réplicas, α puede estimarse empíricamente a partir de los datos. Puesto que la combinación i-ésimo de los tratamientos $\sigma_{y_i} \propto \mu_i^{\alpha} = \theta\mu_i^{\alpha}$, donde θ es una constante de proporcionalidad, pueden tomarse logaritmos para obtener

$$\log \sigma_{y_i} = \log \theta + \alpha \log \mu_i \tag{3-22}$$

Por lo tanto, una gráfica de $\log \sigma_{y_i}$ contra $\log \mu_i$ sería una línea recta con pendiente α. Puesto que no se conocen σ_{y_i} y μ_i, pueden sustituirse estimaciones razonables de ellos en la ecuación 3-22 y usar la pendiente del ajuste de la línea recta resultante como estimación de α. De manera típica, se usaría la desviación estándar S_i y el promedio $\bar{y}_{i.}$ del tratamiento i-ésimo (o, en términos más generales, la combinación i-ésima de los tratamientos o conjunto de condiciones experimentales) para estimar σ_{y_i} y μ_i.

Para investigar la posibilidad de usar una transformación para estabilizar la varianza en los datos de la descarga pico del ejemplo 3-5, en la figura 3-8 se grafica log S_i contra log$\bar{y}_{i.}$. La pendiente de la recta que pasa por estos cuatro puntos está cerca de 1/2 y, por la tabla 3-9, esto implica que la transformación de la raíz cuadrada puede ser apropiada. El análisis de varianza de los datos transformados $y^* = \sqrt{y}$ se presenta en la tabla 3-10, y en la figura 3-9 se muestra una gráfica de los residuales contra la respuesta predicha. Esta gráfica residual muestra una mejoría sensible en comparación con la figura 3-7, por lo que se conclu-

Tabla 3-10 Análisis de varianza de los datos transformados de la descarga pico, $y* = \sqrt{y}$

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0	Valor P
Métodos	32.6842	3	10.8947	76.99	<0.001
Error	2.6884	19	0.1415		
Total	35.3726	22			

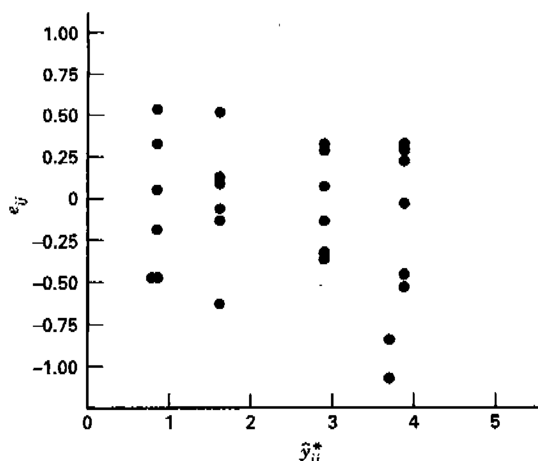


Figura 3-9 Gráfica de los residuales de los datos transformados contra $\hat{y}^*_{ij}$ para los datos de la descarga pico del ejemplo 3-5.

ye que la transformación de la raíz cuadrada ha sido útil. Observe que en la tabla 3-10 se han reducido los grados de libertad del error en 1 para tomar en consideración el uso de los datos para estimar el parámetro de la transformación α.

En la práctica, muchos experimentadores seleccionan la forma de la transformación probando varias alternativas y observando el efecto de cada transformación en la gráfica de los residuales contra la respuesta predicha. Entonces se selecciona la transformación que produjo la gráfica residual más satisfactoria.

3-4.4 Gráficas de los residuales contra otras variables

Si se han recolectado datos de cualesquiera otras variables que posiblemente pudieran afectar la respuesta, los residuales deberán graficarse contra estas variables. Por ejemplo, en el experimento de la resistencia a la tensión del ejemplo 3-1, la resistencia puede ser afectada de manera significativa por el espesor de la fibra, por lo que los residuales deberán graficarse contra el espesor de la fibra. Si se usaron diferentes máquinas de prueba para recolectar los datos, los residuales deberán graficarse contra las máquinas. Los patrones en tales gráficas residuales implican que la variable afecta la respuesta. Esto sugiere que la variable debería controlarse con mayor atención en experimentos futuros o que debería incluirse en el análisis.

3-5 INTERPRETACIÓN PRÁCTICA DE LOS RESULTADOS

Después de realizar el experimento, llevar a cabo el análisis estadístico e investigar los supuestos fundamentales, el experimentador está listo para sacar conclusiones prácticas acerca del problema bajo estudio. Muchas veces esto es relativamente fácil, y ciertamente en los experimentos sencillos que se han considerado hasta este punto, esto podría hacerse de manera un tanto informal, tal vez mediante la inspección de las representaciones gráficas, como los diagramas de caja y el diagrama de dispersión de las figuras 3-1 y 3-2. Sin embargo, en algunos casos es necesario aplicar técnicas más formales. En esta sección se presentarán algunas de ellas.

3-5.1 Un modelo de regresión

Los factores que intervienen en un experimento pueden ser **cuantitativos** o **cualitativos.** Un factor cuantitativo es aquel cuyos niveles pueden asociarse con puntos en una escala numérica, como la temperatura, la presión o el tiempo. Los factores cualitativos, por otra parte, son aquellos cuyos niveles no pueden ordenarse por magnitud. Los operadores, los lotes de materia prima y los cambios de turno son factores cualitativos típicos, ya que no existe ninguna razón para ordenarlos bajo algún criterio numérico particular.

En lo que se refiere al diseño inicial y al análisis del experimento, ambos tipos de factores se tratan de manera idéntica. El experimentador está interesado en determinar las diferencias, en caso de haberlas, entre los niveles de los factores. Si el factor es cualitativo, como los operadores, no tiene sentido considerar la respuesta de una corrida subsecuente en un nivel intermedio del factor. Sin embargo, con un factor cuantitativo como el tiempo, el experimentador tiene interés por lo general en el rango completo de los valores usados, particularmente la respuesta de una corrida subsecuente con un nivel intermedio del factor. Es decir, si se usan los niveles 1.0, 2.0 y 3.0 en el experimento, tal vez quiera predecir la respuesta de 2.5 horas. Por lo tanto, el experimentador tiene con frecuencia interés en desarrollar una ecuación de interpolación para la variable de respuesta del experimento. Esta ecuación es un **modelo empírico** del proceso que se ha estudiado.

Al enfoque general para ajustar modelos empíricos se le llama **análisis de regresión,** el cual se analiza en detalle en el capítulo 10. Véase también el material suplementario del texto para este capítulo. En esta sección se ilustra brevemente la técnica utilizando los datos de la resistencia a la tensión del ejemplo 3-1.

En la figura 3-10 se presenta el diagrama de dispersión de la resistencia a la tensión y contra el peso porcentual del algodón x en la tela para el experimento del ejemplo 3-1. Los círculos huecos de la gráfica son las resistencias de tensión promedio para cada valor x del peso porcentual del algodón. A partir del examen del diagrama de dispersión, es evidente que la relación entre la resistencia a la tensión y el peso

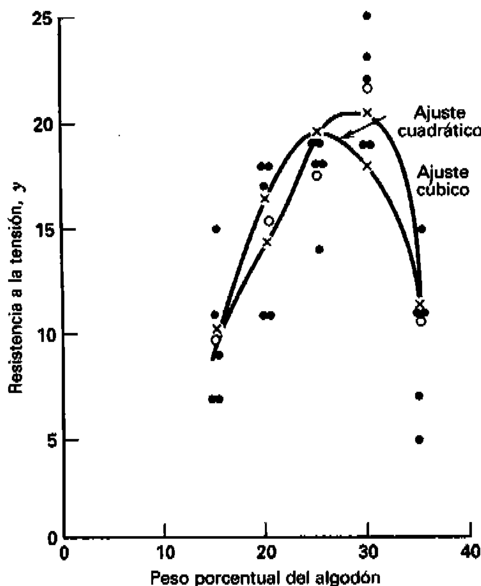


Figura 3-10 Diagrama de dispersión para los datos de la resistencia a la tensión del ejemplo 3-1.

porcentual del algodón no es lineal. Como una primera aproximación, podría intentarse ajustar un **modelo cuadrático** para los datos, por ejemplo

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \varepsilon$$

donde β_0, β_1 y β_2 son parámetros desconocidos que deberán estimarse y ε es un término del error aleatorio. El método que se usa con mayor frecuencia para estimar los parámetros en un modelo como éste es el método de mínimos cuadrados. Éste consiste en elegir estimaciones de las β tales que minimicen la suma de cuadrados de los errores (las ε). El ajuste de mínimos cuadrados en el ejemplo que se considera aquí es

$$\hat{y} = -39.9886 + 4.596x - 0.0886x^2$$

(Si el lector no está familiarizado con los métodos de regresión, vea el capítulo 10 y el material suplementario del texto para este capítulo.)

En la figura 3-10 se muestra este modelo cuadrático. No parece muy satisfactorio, ya que subestima de manera drástica las respuestas para $x = 30\%$ de algodón y sobrestima las respuestas para $x = 25\%$. Quizá pueda lograrse un mejoramiento agregando un término cúbico en x. El ajuste con el **modelo cúbico** resultante es

$$\hat{y} = 62.6114 - 9.0114x + 0.4814x^2 - 0.0076x^3$$

Este ajuste cúbico se ilustra también en la figura 3-10. El modelo cúbico parece mejor que el cuadrático porque proporciona un ajuste mejor para $x = 25$ y $x = 30\%$ de algodón.

En general, sería preferible hacer el ajuste con el polinomio de orden menor que describa adecuadamente el sistema o proceso. En este ejemplo, el polinomio cúbico parece un mejor ajuste que el cuadrático, por lo que la complejidad adicional del modelo cúbico se justifica. Sin embargo, seleccionar el orden del polinomio de aproximación no siempre es fácil, y es relativamente sencillo excederse en el ajuste, es decir, agregar polinomios de orden superior que no mejoran en realidad el ajuste pero que incrementan la complejidad del modelo y con frecuencia demeritan su utilidad como predictor o ecuación de interpolación.

En este ejemplo, el modelo empírico podría usarse para predecir la resistencia a la tensión media para los valores del peso porcentual del algodón dentro de la región de experimentación. En otros casos, el modelo empírico podría usarse para la **optimización del proceso,** es decir, para encontrar los niveles de las variables del diseño que dan como resultado los mejores valores de la respuesta. Más adelante se analizarán e ilustrarán en detalle estos problemas.

3-5.2 Comparaciones entre las medias de los tratamientos

Suponga que al hacer el análisis de varianza en el modelo con efectos fijos se rechaza la hipótesis nula. Por lo tanto, hay diferencias entre las medias de los tratamientos, pero no se especifica exactamente *cuáles* medias difieren. En ocasiones pueden ser de utilidad en esta situación las comparaciones y los análisis adicionales entre **grupos** de las medias de los tratamientos. La media del tratamiento i-ésimo se define como $\mu_i = \mu + \tau_i$, y μ_i se estima con $\bar{y}_{i.}$. Las comparaciones entre las medias de los tratamientos se hacen ya sea en términos de los totales de los tratamientos $\{y_{i.}\}$ o bien de los promedios de los tratamientos $\{\bar{y}_{i.}\}$. Es común llamar a los procedimientos para hacer estas comparaciones **métodos de comparaciones múltiples.** En varias de las secciones siguientes se analizan los métodos para hacer comparaciones entre las medias de los tratamientos individuales o de grupos de estas medias.

3-5.3 Comparaciones gráficas de medias

Es muy sencillo desarrollar un procedimiento gráfico para la comparación de las medias después de un análisis de varianza. Suponga que el factor de interés tiene a niveles y que $\bar{y}_{1.}, \bar{y}_{2.}, ..., \bar{y}_{a.}$ son los promedios de los tratamientos. Si se conoce σ, el promedio de cualquier tratamiento tendría una desviación estándar $\sigma/\sqrt{n}$. Por consiguiente, si todas las medias de los niveles del factor son idénticas, las medias muestrales observadas $\bar{y}_{i.}$ se comportarían como un conjunto de observaciones tomadas al azar de una distribución normal con media $\bar{y}_{..}$ y desviación estándar $\sigma/\sqrt{n}$. Visualice una distribución normal con la capacidad de ser deslizada sobre un eje abajo del cual están graficadas $\bar{y}_{1.}, \bar{y}_{2.}, ..., \bar{y}_{a.}$. Si todas las medias de los tratamientos son iguales, deberá haber una posición de esta distribución que haga evidente que los valores $\bar{y}_{i.}$ se sacaron de la misma distribución. Si no es éste el caso, los valores $\bar{y}_{i.}$ que *no* parecen haberse sacado de esta distribución se asocian con los niveles del factor que producen respuestas medias diferentes.

La única falla en esta lógica es que σ es desconocida. Sin embargo, puede sustituirse σ con $\sqrt{MS_E}$ del análisis de varianza y usar una distribución t con un factor de escalación $\sqrt{MS_E / n}$ en lugar de la distribución normal. En la figura 3-11 se muestra este arreglo para los datos de la resistencia a la tensión del ejemplo 3-1.

Para trazar la distribución t de la figura 3-11, simplemente se multiplica el valor de la abscisa t por el factor de escalación

$$\sqrt{MS_E / n} = \sqrt{8.06 / 5} = 1.27$$

y se grafica contra la ordenada de t en ese punto. Dada la gran similitud entre la distribución t y la normal, excepto porque la primera es un poco más plana cerca del centro y tiene colas más largas, por lo general este trazo se construye fácilmente a ojo. Si quiere obtenerse mayor precisión, en Box, Hunter y Hunter [18] se incluye una tabla de los valores de las abscisas t y las ordenadas correspondientes. La distribución puede tener un origen arbitrario, aun cuando por lo general es mejor uno que esté en la región de los valores $\bar{y}_{i.}$ que van a compararse. En la figura 3-11, el origen es 15 lb/pulg^2.

Visualice ahora el desplazamiento de la distribución t de la figura 3-11 sobre el eje horizontal y examine las cinco medias graficadas en la figura. Observe que no hay ninguna posición de la distribución tal que los cinco promedios puedan considerarse como observaciones típicas seleccionadas al azar de la distribución. Esto implica que las cinco medias no son iguales; por lo tanto, la figura es una representación gráfica de los resultados del análisis de varianza. La figura indica que 30% de algodón produce resistencias a la tensión mucho más altas que 20 o 25 por ciento de algodón (las cuales son aproximadamente iguales), y que 15 o 35 por ciento de algodón (las cuales son aproximadamente iguales) producirían resistencias a la tensión aún más bajas.

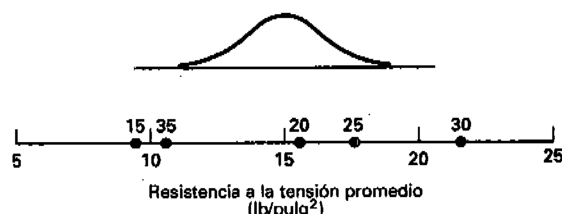


Figura 3-11 Promedio de la resistencia a la tensión del experimento del peso porcentual del algodón en relación con una distribución t con un factor de escalación $\sqrt{MS_E / n} = \sqrt{8.06 / 5} = 1.27$.

Este procedimiento simple es una técnica aproximada pero eficaz en muchos problemas de comparaciones múltiples. Sin embargo, existen métodos más formales. A continuación se presenta una breve revisión de algunos de estos procedimientos.

3-5.4 Contrastes

Muchos métodos de comparaciones múltiples utilizan el concepto de contraste. Considere el problema de la prueba de la fibra sintética del ejemplo 3-1. Puesto que se rechazó la hipótesis nula, se sabe que algunos pesos porcentuales del algodón producen resistencias a la tensión diferentes que otros, pero, ¿cuáles son los que causan en realidad esta diferencia? Al principio del experimento podría sospecharse que los niveles 4 y 5 del peso porcentual del algodón (30 y 35 por ciento) producen la misma resistencia a la tensión, lo cual implicaría que la hipótesis por probar sería

$$H_0: \mu_4 = \mu_5$$
$$H_1: \mu_4 \neq \mu_5$$

o, de manera equivalente,

$$\begin{aligned} H_0&: \mu_4 - \mu_5 = 0 \\ H_1&: \mu_4 - \mu_5 \neq 0 \end{aligned} \tag{3-23}$$

Si desde el principio del experimento se hubiera sospechado que el *promedio* de los niveles más bajos del peso porcentual del algodón (1 y 2) no difería del *promedio* de los niveles más altos del peso porcentual del algodón (4 y 5), entonces la hipótesis habría sido

$$H_0: \mu_1 + \mu_2 = \mu_4 + \mu_5$$
$$H_1: \mu_1 + \mu_2 \neq \mu_4 + \mu_5$$

o

$$\begin{aligned} H_0&: \mu_1 + \mu_2 - \mu_4 - \mu_5 = 0 \\ H_1&: \mu_1 + \mu_2 - \mu_4 - \mu_5 = 0 \end{aligned} \tag{3-24}$$

En general, un **contraste** es una combinación lineal de parámetros de la forma

$$\Gamma = \sum_{i=1}^{a} c_i \mu_i$$

donde las **constantes de los contrastes** $c_1, c_2, ..., c_a$ suman cero; es decir, $\Sigma_{i=1}^{a} c_i = 0$. Las dos hipótesis anteriores pueden expresarse en términos de contrastes:

$$\begin{aligned} H_0&: \sum_{i=1}^{a} c_i \mu_i = 0 \\ H_1&: \sum_{i=1}^{a} c_i \mu_i \neq 0 \end{aligned} \tag{3-25}$$

Las constantes de los contrastes para las hipótesis de la ecuación 3-23 son $c_1 = c_2 = c_3 = 0$, $c_4 = +1$ y $c_5 = -1$, mientras que para las hipótesis de la ecuación 3-24 son $c_1 = c_2 = +1$, $c_3 = 0$, y $c_4 = c_5 = -1$.

Las pruebas de hipótesis que incluyen contrastes pueden hacerse de dos maneras básicas. En el primer método se utiliza la prueba t. El contraste de interés se escribe en términos de los **totales de los tratamientos,** obteniéndose

$$C = \sum_{i=1}^{a} c_i y_{i.}$$

La varianza de C es

$$V(C) = n\sigma^2 \sum_{i=1}^{a} c_i^2 \tag{3-26}$$

cuando los tamaños de las muestras de cada tratamiento son iguales. Si la hipótesis nula de la ecuación 3-25 es verdadera, el cociente

$$\frac{\sum_{i=1}^{a} c_i y_{i.}}{\sqrt{n\sigma^2 \sum_{i=1}^{a} c_i^2}}$$

tiene la distribución $N(0, 1)$. Entonces se sustituiría la varianza desconocida σ^2 con su estimación, el error cuadrático medio MS_E, y se utilizaría el estadístico

$$t_0 = \frac{\sum_{i=1}^{a} c_i y_{i.}}{\sqrt{nMS_E \sum_{i=1}^{a} c_i^2}} \tag{3-27}$$

para probar las hipótesis de la ecuación 3-25. La hipótesis nula se rechazaría si $|t_0|$ de la ecuación 3-27 excede $t_{\alpha/2,\, N-a}$.

En el segundo enfoque se utiliza la prueba F. Entonces, el cuadrado de una variable aleatoria t con ν grados de libertad es una variable aleatoria F con un grado de libertad en el numerador y ν grados de libertad en el denominador. Por lo tanto, puede obtenerse

$$F_0 = t_0^2 = \frac{\left(\sum_{i=1}^{a} c_i y_{i.}\right)^2}{nMS_E \sum_{i=1}^{a} c_i^2} \tag{3-28}$$

como un estadístico F para probar la ecuación 3-25. La hipótesis nula se rechazaría si $F_0 > F_{\alpha,1,N-a}$. Este estadístico de prueba de la ecuación 3-28 puede escribirse como

$$F_0 = \frac{MS_C}{MS_E} = \frac{SS_C / 1}{MS_E}$$

donde la suma de cuadrados de los contrastes con un solo grado de libertad es

$$SS_C = \frac{\left(\sum_{i=1}^{a} c_i y_{i.}\right)^2}{n\sum_{i=1}^{a} c_i^2} \tag{3-29}$$

Intervalo de confianza para un contraste

En lugar de probar hipótesis acerca de un contraste, puede ser más útil construir un intervalo de confianza. Entonces el contraste suele expresarse en términos de los promedios de los tratamientos $\bar{y}_{i.}$ Suponga que el contraste de interés es

$$\Gamma = \sum_{i=1}^{a} c_i \mu_i$$

Al sustituir las medias de los tratamientos con los promedios de los tratamientos se obtiene

$$C = \sum_{i=1}^{a} c_i \bar{y}_{i.}$$

y

$$V(C) = \frac{\sigma^2}{n} \sum_{i=1}^{a} c_i^2$$

cuando los tamaños de las muestras son iguales. Si se usa MS_E para estimar σ^2, el intervalo de confianza de 100(1 − α) por ciento para el contraste $\Sigma_{i=1}^{a} c_i \mu_i$ es

$$\sum_{i=1}^{a} c_i \bar{y}_{i.} - t_{\alpha/2,N-a} \sqrt{\frac{MS_E}{n} \sum_{i=1}^{a} c_i^2} \leq \sum_{i=1}^{a} c_i \mu_i \leq \sum_{i=1}^{a} c_i \bar{y}_{i.} + t_{\alpha/2,N-a} \sqrt{\frac{MS_E}{n} \sum_{i=1}^{a} c_i^2} \qquad (3\text{-}30)$$

Evidentemente, si este intervalo de confianza incluye al cero, no podría rechazarse la hipótesis nula en la ecuación 3-25.

Contraste estandarizado

Cuando hay interés en más de un contraste, con frecuencia es útil evaluarlos en la misma escala. Una forma de hacer esto es estandarizando el contraste para que su varianza sea σ^2. Si el contraste $\Sigma_{i=1}^{a} c_i \mu_i$ se expresa en términos de los totales de los tratamientos como $\Sigma_{i=1}^{a} c_i y_{i.}$, al dividirlo por $\sqrt{n\Sigma_{i=1}^{a} c_i^2}$ se obtendrá un contraste estandarizado con varianza σ^2. Entonces el **contraste estandarizado** es en realidad

$$\sum_{i=1}^{a} c_i^* y_{i.}$$

donde

$$c_i^* = \frac{c_i}{\sqrt{n \sum_{i=1}^{a} c_i^2}}$$

Tamaños de las muestras desiguales

Cuando los tamaños de las muestras de cada tratamiento son diferentes, se introducen modificaciones menores en los resultados anteriores. Primero, observe que la definición de un contraste requiere ahora que

$$\sum_{i=1}^{a} n_i c_i = 0$$

Otros cambios requeridos son directos. Por ejemplo, el estadístico t de la ecuación 3-27 queda como

$$t_0 = \frac{\sum_{i=1}^{a} c_i y_{i.}}{\sqrt{MS_E \sum_{i=1}^{a} n_i c_i^2}}$$

y la suma de cuadrados de los contrastes de la ecuación 3-29 queda como

$$SS_C = \frac{\left(\sum_{i=1}^{a} c_i y_{i.}\right)^2}{\sum_{i=1}^{a} n_i c_i^2}$$

3-5.5 Contrastes ortogonales

Un caso especial útil del procedimiento de la sección 3-5.4 es el de los **contrastes ortogonales.** Dos contrastes con coeficientes $\{c_i\}$ y $\{d_i\}$ son ortogonales si

$$\sum_{i=1}^{a} c_i d_i = 0$$

o, para un diseño no balanceado, si

$$\sum_{i=1}^{a} n_i c_i d_i = 0$$

Para a tratamientos, el conjunto de $a-1$ contrastes ortogonales hace la partición de la suma de cuadrados debida a los tratamientos en $a-1$ componentes independientes con un solo grado de libertad. Por lo tanto, las pruebas que se realizan en los contrastes ortogonales son independientes.

Existen varias maneras de elegir los coeficientes de los contrastes ortogonales para un conjunto de tratamientos. En general, algún elemento en la naturaleza del experimento deberá sugerir las comparaciones que son de interés. Por ejemplo, si hay $a = 3$ tratamientos, donde el tratamiento 1 es de control y donde los niveles del factor en los tratamientos 2 y 3 son de interés para el experimentador, los contrastes ortogonales apropiados podrían ser los siguientes:

Tratamiento	Coeficientes de los contrastes ortogonales	
1 (control)	–2	0
2 (nivel 1)	1	–1
3 (nivel 2)	1	1

Observe que el contraste 1 con $c_i = -2, 1, 1$ compara el efecto promedio del factor con el control, mientras que el contraste 2 con $d_i = 0, -1, 1$ compara los dos niveles del factor de interés.

En general, el método de contrastes (o de contrastes ortogonales) es útil para lo que se llama **comparaciones preplaneadas.** Es decir, los contrastes se especifican antes de llevar a cabo el experimento y de examinar los datos. La razón de esto es que, si las comparaciones se seleccionan después de examinar los

datos, la mayoría de los experimentadores construirían pruebas que corresponderían con las diferencias grandes observadas en las medias. Estas diferencias grandes podrían ser el resultado de la presencia de efectos reales o podrían ser el resultado del error aleatorio. Si los experimentadores se inclinan consistentemente a escoger las diferencias más grandes para hacer las comparaciones, inflarán el error tipo I de la prueba porque es probable que, en un porcentaje inusualmente elevado de las comparaciones seleccionadas, las diferencias observadas serán el resultado del error. Al examen de los datos para seleccionar las comparaciones de interés potencial se le llama con frecuencia **curioseo o sondeo de datos.** El método de Scheffé para todas las comparaciones, el cual se comenta en la sección siguiente, permite el curioseo o sondeo de datos.

EJEMPLO 3-6

Considere los datos del ejemplo 3-1. Hay cinco medias de los tratamientos y cuatro grados de libertad entre estos tratamientos. Suponga que antes de correr el experimento se especificó la siguiente serie de comparaciones entre las medias de los tratamientos (y sus contrastes asociados):

Hipótesis	Contraste
$H_0{:}\mu_4 = \mu_5$	$C_1 = -y_{4.} + y_{5.}$
$H_0{:}\mu_1 + \mu_3 = \mu_4 + \mu_5$	$C_2 = y_{1.} + y_{3.} - y_{4.} - y_{5.}$
$H_0{:}\mu_1 = \mu_3$	$C_3 = y_{1.} - y_{3.}$
$H_0{:}4\mu_2 = \mu_1 + \mu_3 + \mu_4 + \mu_5$	$C_4 = -y_{1.} + 4y_{2.} - y_{3.} - y_{4.} - y_{5.}$

Observe que los coeficientes de los contrastes son ortogonales. Utilizando los datos de la tabla 3-4, se encuentra que los valores numéricos de los contrastes y de las sumas de cuadrados son los siguientes:

$$C_1 = -1(108)+1(54) = -54 \qquad SS_{C_1} = \frac{(-54)^2}{5(2)} = 291.60$$

$$C_2 = +1(49) +1(88)-1(108)-1(54) = -25 \qquad SS_{C_2} = \frac{(-25)^2}{5(4)} = 31.25$$

$$C_3 = +1(49) -1(88) = -39 \qquad SS_{C_3} = \frac{(-39)^2}{5(2)} = 152.10$$

$$C_4 = -1(49)+4(77)-1(88)-1(108)-1(54) = 9 \qquad SS_{C_4} = \frac{(9)^2}{5(20)} = 0.81$$

Tabla 3-11 Análisis de varianza de los datos de la resistencia a la tensión

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0	Valor P
Peso porcentual del algodón	475.76	4	118.94	14.76	<0.001
contrastes ortogonales					
$C_1{:}\mu_4 = \mu_5$	(291.60)	1	291.60	36.18	<0.001
$C_2{:}\mu_1 + \mu_3 = \mu_4 + \mu_5$	(31.25)	1	31.25	3.88	0.06
$C_3{:}\mu_1 = \mu_3$	(152.10)	1	152.10	18.87	<0.001
$C_4{:}4\mu_2 = \mu_1 + \mu_3 + \mu_4 + \mu_5$	(0.81)	1	0.81	0.10	0.76
Error	161.20	20	8.06		
Total	636.96	24			

Estas sumas de cuadrados de los contrastes hacen la partición completa de la suma de cuadrados de los tratamientos. Las pruebas de estos contrastes ortogonales se incorporan por lo general en el análisis de varianza, como se muestra en la tabla 3-11. Por los valores P se concluye que hay diferencias significativas entre los niveles 4 y 5 y 1 y 3 del peso porcentual del algodón, pero que el *promedio* de los niveles 1 y 3 no difiere del promedio de los niveles 4 y 5 con el nivel $\alpha = 0.05$, y que el nivel 2 no difiere del promedio de los otros cuatro niveles.

3-5.6 Método de Scheffé para comparar todos los contrastes

En muchas situaciones, los experimentadores pueden no conocer de antemano cuáles son los contrastes que quieren comparar, o pueden tener interés en más de $a - 1$ posibles comparaciones. En muchos experimentos de exploración, las comparaciones de interés sólo se descubren después del examen preliminar de los datos. Scheffé [98a] ha propuesto un método para comparar todos y cada uno de los contrastes posibles entre las medias de los tratamientos. En el método de Scheffé, el error tipo I es a lo sumo α para cualquiera de las comparaciones posibles.

Suponga que se ha determinado un conjunto de m contrastes

$$\Gamma_u = c_{1u}\mu_1 + c_{2u}\mu_2 + \cdots + c_{au}\mu_a \qquad u = 1, 2, \ldots, m \tag{3-31}$$

en las medias de los tratamientos de interés. El contraste correspondiente de los promedios de los tratamientos $\bar{y}_{i.}$ es

$$C_u = c_{1u}\bar{y}_{1.} + c_{2u}\bar{y}_{2.} + \cdots + c_{au}\bar{y}_{a.} \qquad u = 1, 2, \ldots, m \tag{3-32}$$

y el **error estándar** de este contraste es

$$S_{C_u} = \sqrt{MS_E \sum_{i=1}^{a} (c_{iu}^2 / n_i)} \tag{3-33}$$

donde n_i es el número de observaciones en el tratamiento i-ésimo. Puede demostrarse que el valor crítico contra el que deberá compararse C_u es

$$S_{\alpha,u} = S_{C_u}\sqrt{(a-1)F_{\alpha,a-1,N-a}} \tag{3-34}$$

Para probar la hipótesis de que el contraste Γ_u difiere de manera significativa de cero, se compara C_u con el valor crítico. Si $|C_u| > S_{\alpha,u}$, se rechaza la hipótesis de que el contraste Γ_u es igual a cero.

El procedimiento de Scheffé puede usarse también para formar intervalos de confianza para todos los contrastes posibles entre las medias de los tratamientos. Los intervalos resultantes, por ejemplo $C_u - S_{\alpha,u} \leq \Gamma_u \leq C_u + S_{\alpha,u}$, son **intervalos de confianza simultáneos** por cuanto la probabilidad de que todos ellos sean verdaderos simultáneamente es al menos $1 - \alpha$.

Para ilustrar el procedimiento, considere los datos del ejemplo 3-1 y suponga que los contrastes de interés son

$$\Gamma_1 = \mu_1 + \mu_3 - \mu_4 - \mu_5$$

y

$$\Gamma_2 = \mu_1 - \mu_4$$

Los valores numéricos de estos contrastes son

$$\begin{aligned} C_1 &= \bar{y}_{1.} + \bar{y}_{3.} - \bar{y}_{4.} - \bar{y}_{5.} \\ &= 9.80 + 17.60 - 21.60 - 10.80 \\ &= 5.00 \end{aligned}$$

y

$$\begin{aligned} C_2 &= \bar{y}_{1.} - \bar{y}_{4.} \\ &= 9.80 - 21.60 \\ &= -11.80 \end{aligned}$$

y los errores estándar se encuentran con la ecuación 3-33 como

$$S_{C_1} = \sqrt{MS_E \sum_{i=1}^{5} (c_{i1}^2 / n_i)} = \sqrt{8.06(1+1+1+1)/5} = 2.54$$

y

$$S_{C_2} = \sqrt{MS_E \sum_{i=1}^{5} (c_{i2}^2 / n_i)} = \sqrt{8.06(1+1)/5} = 1.80$$

Por la ecuación 3-34, los valores críticos de 1% son

$$S_{0.01,1} = S_{C_1} \sqrt{(a-1)F_{0.01,a-1,N-a}} = 2.54\sqrt{4(4.43)} = 10.69$$

y

$$S_{0.01,2} = S_{C_2} \sqrt{(a-1)F_{0.01,a-1,N-a}} = 1.80\sqrt{4(4.43)} = 7.58$$

Puesto que $|C_1| < S_{0.01,1}$, se concluye que el contraste $\Gamma_1 = \mu_1 + \mu_3 - \mu_4 - \mu_5$ es igual a cero; es decir, no existe evidencia sólida para concluir que las medias de los tratamientos 1 y 3 como grupo difieren de las medias de los tratamientos 4 y 5 como grupo. Sin embargo, como $|C_2| > S_{0.01,2}$, se concluye que el contraste $\Gamma_2 = \mu_1 - \mu_4$ no es igual a cero; es decir, las resistencias medias de los tratamientos 1 y 4 difieren significativamente.

En muchas situaciones prácticas, querrán compararse sólo **pares de medias.** Frecuentemente, es posible determinar cuáles son las medias que difieren probando las diferencias entre *todos* los pares de medias de los tratamientos. Por lo tanto, el interés se encuentra en los contrastes de la forma $\Gamma = \mu_i - \mu_j$ para toda $i \neq j$. Aun cuando el método de Scheffé podría aplicarse fácilmente a este problema, no es el procedimiento más sensible para tales comparaciones. Se pasa ahora a la consideración de los métodos diseñados específicamente para las comparaciones por pares entre todas las a medias poblacionales.

3-5.7 Comparación de pares de medias de tratamientos

Suponga que el interés se encuentra en comparar todos los pares de a medias de tratamientos y que las hipótesis nulas que quieren probarse son $H_0{:}\mu_i = \mu_j$ para toda $i \neq j$. A continuación se presentan cuatro métodos para hacer estas comparaciones.

Prueba de Tukey

Suponga que, después de un análisis de varianza en el que se ha rechazado la hipótesis nula de la igualdad de las medias de los tratamientos, quieren probarse todas las comparaciones de las medias por pares:

$$\begin{aligned} H_0&: \mu_i = \mu_j \\ H_1&: \mu_i \neq \mu_j \end{aligned}$$

para toda $i \neq j$. Tukey [111d] propuso un procedimiento para probar hipótesis para las que el nivel de significación global es exactamente α cuando los tamaños de las muestras son iguales y es a lo sumo α cuando

los tamaños de las muestras no son iguales. Este procedimiento puede usarse también para contraer los intervalos de confianza para las diferencias en todos los pares de medias. Para estos intervalos, el nivel de confianza simultáneo es de 100(1 – α) por ciento cuando los tamaños de las muestras son iguales y de al menos 100(1 – α) por ciento cuando los tamaños de las muestras no son iguales. Se trata de un procedimiento excelente para curiosear sobre los datos cuando el interés se centra en pares de medias.

El procedimiento de Tukey hace uso de la distribución del **estadístico del rango studentizado**

$$q = \frac{\bar{y}_{\text{máx}} - \bar{y}_{\text{mín}}}{\sqrt{MS_E / n}}$$

donde $\bar{y}_{\text{máx}}$ y $\bar{y}_{\text{mín}}$ son las medias muestrales mayor y menor, respectivamente, sacadas de un grupo de p medias muestrales. La tabla VIII del apéndice contiene los valores de $q_\alpha(p, f)$, los puntos porcentuales α superiores de q, donde f es el número de grados de libertad asociados con MS_E. Para tamaños de las muestras iguales, la prueba de Tukey declara que dos medias son significativamente diferentes si el valor absoluto de sus diferencias muestrales excede

$$T_\alpha = q_\alpha(a, f)\sqrt{\frac{MS_E}{n}} \tag{3-35}$$

De manera equivalente, podría construirse una serie de intervalos de confianza de 100(1 – α) por ciento para todos los pares de medias de la siguiente manera:

$$\bar{y}_{i.} - \bar{y}_{j.} - q_\alpha(a, f)\sqrt{\frac{MS_E}{n}} \leq \mu_i - \mu_j \leq \bar{y}_{i.} - \bar{y}_{j.} + q_\alpha(a, f)\sqrt{\frac{MS_E}{n}}, \quad i \neq j. \tag{3-36}$$

Cuando los tamaños de las muestras no son iguales, las ecuaciones 3-35 y 3-36 quedan como

$$T_\alpha = \frac{q_\alpha(a, f)}{\sqrt{2}}\sqrt{MS_E\left(\frac{1}{n_i} + \frac{1}{n_j}\right)} \tag{3-37}$$

y

$$\bar{y}_{i.} - \bar{y}_{j.} - \frac{q_\alpha(a, f)}{\sqrt{2}}\sqrt{MS_E\left(\frac{1}{n_i} + \frac{1}{n_j}\right)} \leq \mu_i - \mu_j \leq \bar{y}_{i.} - \bar{y}_{j.} + \frac{q_\alpha(a, f)}{\sqrt{2}}\sqrt{MS_E\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}, i \neq j \tag{3-38}$$

respectivamente. A la versión para tamaños de las muestras diferentes se le llama en ocasiones el **procedimiento Tukey-Kramer.**

EJEMPLO 3-7 ···

Para ilustrar la prueba de Tukey, se usan los datos del experimento del peso porcentual del algodón del ejemplo 3-1. Con $\alpha = 0.05$ y $f = 20$ grados de libertad para el error, en la tabla VIII del apéndice se obtiene $q_{0.05}(5, 20) = 4.23$. Por lo tanto, por la ecuación 3-35,

$$T_{0.05} = q_{0.05}(5, 20)\sqrt{\frac{MS_E}{n}} = 4.23\sqrt{\frac{8.06}{5}} = 5.37$$

$\bar{y}_{1.}$	$\bar{y}_{5.}$	$\bar{y}_{2.}$	$\bar{y}_{3.}$	$\bar{y}_{4.}$
9.8	10.8	15.4	17.6	21.6

Figura 3-12 Resultados de la prueba de Tukey.

Por lo tanto, cualquier par de promedios de los tratamientos que difieran en valor absoluto por más de 5.37 implicaría que el par correspondiente de medias poblacionales son significativamente diferentes. Los cinco promedios de los tratamientos son

$$\bar{y}_{1.} = 9.8 \qquad \bar{y}_{2.} = 15.4 \qquad \bar{y}_{3.} = 17.6 \qquad \bar{y}_{4.} = 21.6 \qquad \bar{y}_{5.} = 10.8$$

y las diferencias en los promedios son

$$\begin{aligned}
\bar{y}_{1.} - \bar{y}_{2.} &= 9.8 - 15.4 = -5.6^{*} \\
\bar{y}_{1.} - \bar{y}_{3.} &= 9.8 - 17.6 = -7.8^{*} \\
\bar{y}_{1.} - \bar{y}_{4.} &= 9.8 - 21.6 = -11.8^{*} \\
\bar{y}_{1.} - \bar{y}_{5.} &= 9.8 - 10.8 = -1.0 \\
\bar{y}_{2.} - \bar{y}_{3.} &= 15.4 - 17.6 = -2.2 \\
\bar{y}_{2.} - \bar{y}_{4.} &= 15.4 - 21.6 = -6.2^{*} \\
\bar{y}_{2.} - \bar{y}_{5.} &= 15.4 - 10.8 = 4.6 \\
\bar{y}_{3.} - \bar{y}_{4.} &= 17.6 - 21.6 = -4.0 \\
\bar{y}_{3.} - \bar{y}_{5.} &= 17.6 - 10.8 = 6.8^{*} \\
\bar{y}_{4.} - \bar{y}_{5.} &= 21.6 - 10.8 = 10.8^{*}
\end{aligned}$$

Los valores marcados con asterisco indican pares de medias que son significativamente diferentes. Suele ser útil trazar una gráfica, como la de la figura 3-12, donde se subraya a los pares de medias que no difieren significativamente. Esta gráfica da una indicación de que las medias de los tratamientos forman tres grupos: μ_1 y μ_5, μ_2 y μ_3, y μ_4. Sin embargo, la pertenencia a estos grupos no es del todo clara.

Cuando se utiliza cualquiera de los procedimientos para probar las medias por pares, ocasionalmente se encuentra que la prueba F global del análisis de varianza es significativa, pero la comparación de las medias por pares falla para revelar cualquier diferencia significativa. Esta situación ocurre porque la prueba F considera simultáneamente todos los contrastes posibles en los que intervienen las medias de los tratamientos, no sólo las comparaciones por pares. Es decir, en los datos a la mano, quizá no todos los contrastes significativos sean de la forma $\mu_i - \mu_j$.

Algunos paquetes de software de computadora presentan comparaciones por pares con intervalos de confianza. Para el procedimiento de Tukey, estos intervalos se calcularían con la ecuación 3-36 o la 3-38, dependiendo de si los tamaños de las muestras son iguales o no.

La deducción del intervalo de confianza de Tukey de la ecuación 3-36 para tamaños de las muestras iguales es directa. Para el estadístico del rango studentizado q se tiene

$$P\left(\frac{\text{máx}(\bar{y}_{i.} - \mu_i) - \text{mín}(\bar{y}_{i.} - \mu_i)}{\sqrt{MS_E / n}} \leq q_\alpha(a, f) \right) = 1 - \alpha$$

Si máx$(\bar{y}_{i.}-\mu_i)$ – mín$(\bar{y}_{i.}-\mu_i)$ es menor o igual que $q_\alpha(a,f)\sqrt{MS_E/n}$, debe ser verdadero que $|(\bar{y}_{i.}-\mu_i)-(\bar{y}_{j.}-\mu_j)| \leq q_\alpha(a,f)\sqrt{MS_E/n}$ para cada par de medias. Por lo tanto,

$$P\left(-q_\alpha(a,f)\sqrt{\frac{MS_E}{n}} \leq \bar{y}_{i.}-\bar{y}_{j.}-(\mu_i-\mu_j)\leq q_\alpha(a,f)\sqrt{\frac{MS_E}{n}}\right)=1-\alpha$$

Al reordenar esta expresión para aislar $\mu_i-\mu_j$ entre las desigualdades se llegará al conjunto de intervalos de confianza simultáneos de $100(1-\alpha)$ por ciento dado en la ecuación 3-38.

El método de la diferencia significativa mínima (LSD) de Fisher

En este procedimiento se utiliza el estadístico F para probar $H_0{:}\mu_i = \mu_j$

$$t_0=\frac{\bar{y}_{i.}-\bar{y}_{j.}}{\sqrt{MS_E\left(\frac{1}{n_i}+\frac{1}{n_j}\right)}} \tag{3-39}$$

Suponiendo una hipótesis alternativa de dos colas, los pares de medias μ_i y μ_j se declararían significativamente diferentes si $|\bar{y}_{i.}-\bar{y}_{j.}| > t_{\alpha/2,N-a}\sqrt{MS_E(1/n_i+1/n_j)}$. A la cantidad

$$\text{LSD} = t_{\alpha/2,N-a}\sqrt{MS_E\left(\frac{1}{n_i}+\frac{1}{n_j}\right)} \tag{3-40}$$

se le llama **diferencia significativa mínima.** Si el diseño es balanceado, $n_1 = n_2 = \cdots = n_a = n$, y

$$\text{LSD} = t_{\alpha/2,N-a}\sqrt{\frac{2MS_E}{n}} \tag{3-41}$$

Para usar el procedimiento LSD de Fisher, simplemente se compara la diferencia observada entre cada par de promedios con la LSD correspondiente. Si $|\bar{y}_{i.}-\bar{y}_{j.}| >$ LSD, se concluye que las medias poblacionales μ_i y μ_j difieren.

EJEMPLO 3-8 ···

Para ilustrar el procedimiento, si se usan los datos del experimento del ejemplo 3-1, la LSD con $\alpha = 0.05$ es

$$\text{LSD} = t_{.025,20}\sqrt{\frac{2MS_E}{n}} = 2.086\sqrt{\frac{2(8.06)}{5}} = 3.75$$

Por lo tanto, cualquier par de promedios de los tratamientos que difiera del valor absoluto por más de 3.75 implicaría que el par correspondiente de medias poblacionales es significativamente diferente. Las diferencias en los promedios son

$$\begin{aligned}
\bar{y}_{1.}-\bar{y}_{2.} &= 9.8-15.4 = -5.6^*\\
\bar{y}_{1.}-\bar{y}_{3.} &= 9.8-17.6 = -7.8^*\\
\bar{y}_{1.}-\bar{y}_{4.} &= 9.8-21.6 = -11.8^*\\
\bar{y}_{1.}-\bar{y}_{5.} &= 9.8-10.8 = -1.0\\
\bar{y}_{2.}-\bar{y}_{3.} &= 15.4-17.6 = -2.2\\
\bar{y}_{2.}-\bar{y}_{4.} &= 15.4-21.6 = -6.2^*\\
\bar{y}_{2.}-\bar{y}_{5.} &= 15.4-10.8 = 4.6^*\\
\bar{y}_{3.}-\bar{y}_{4.} &= 17.6-21.6 = -4.0^*\\
\bar{y}_{3.}-\bar{y}_{5.} &= 17.6-10.8 = 6.8^*\\
\bar{y}_{4.}-\bar{y}_{5.} &= 21.6-10.8 = 10.8^*
\end{aligned}$$

$\bar{y}_{1.}$	$\bar{y}_{5.}$	$\bar{y}_{2.}$	$\bar{y}_{3.}$	$\bar{y}_{4.}$
9.8	10.8	15.4	17.6	21.6

Figura 3-13 Resultados del procedimiento LSD.

Los valores marcados con asterisco indican pares de medias que son significativamente diferentes. En la figura 3-13 se resumen los resultados. Evidentemente, los únicos pares de medias que no difieren significativamente son 1 y 5 y 2 y 3, y el tratamiento 4 produce una resistencia a la tensión significativamente mayor que los otros tratamientos.

Observe que el riesgo global α puede inflarse de manera considerable al utilizar este método. Específicamente, cuando a se hace más grande, el error tipo I del experimento (el cociente del número de experimentos en los que se comete al menos un error tipo I y el número total de experimentos) se hace grande.

Prueba del rango múltiple de Duncan

Un procedimiento muy utilizado para comparar todos los pares de medias es la **prueba del rango múltiple** desarrollada por Duncan [41]. Para aplicar la prueba del rango múltiple de Duncan cuando los tamaños de las muestras son iguales, los a promedios de los tratamientos se arreglan en orden ascendente, y el error estándar de cada promedio se determina como

$$S_{\bar{y}_{i.}} = \sqrt{\frac{MS_E}{n}} \tag{3-42}$$

Para tamaños de las muestras desiguales, se sustituye n en la ecuación 3-42 con la media armónica n_h del $\{n_i\}$, donde

$$n_h = \frac{a}{\sum_{i=1}^{a} (1/n_i)} \tag{3-43}$$

Observe que si $n_1 = n_2 = \cdots = n_a$, $n_h = n$. En la tabla de Duncan de los rangos significativos (tabla VII del apéndice) se obtienen los valores $r_\alpha(p,f)$ para $p = 2, 3, ..., a$, donde α es el nivel de significación y f es el número de grados de libertad del error. Estos rangos se convierten en un conjunto de $a - 1$ rangos mínimos de significación (por ejemplo, R_p) para $p = 2, 3, ..., a$ calculando

$$R_p = r_\alpha(p, f)S_{\bar{y}_{i.}} \qquad \text{para } p = 2, 3, \ldots, a \tag{3-44}$$

Entonces, se prueban las diferencias observadas entre las medias, empezando con la más grande contra la menor, la cual se compararía con el rango mínimo de significación R_a. Después se calcula la diferencia de la mayor y la segunda menor y se compara con el rango mínimo de significación R_{a-1}. Estas comparaciones se continúan hasta que todas las medias se han comparado con la media mayor. Por último, se calcula la diferencia entre la segunda media mayor y la menor y se compara con el rango mínimo de significación R_{a-1}. Este proceso se continúa hasta que se han considerado las diferencias entre todos los $a(a-1)/2$ pares de medias posibles. Si una diferencia observada es mayor que el rango de significación mínima correspondiente, se concluye que el par de medias en cuestión es significativamente diferente. Para evitar

contradicciones, ninguna de las diferencias entre un par de medias se considera significativa si las dos medias en cuestión se localizan entre otras dos medias que no difieren significativamente.

EJEMPLO 3-9 ....

La prueba del rango múltiple de Duncan puede aplicarse al experimento del ejemplo 3-1. Recuerde que $MS_E = 8.06$, $N = 25$, $n = 5$, y hay 20 grados de libertad del error. Al arreglar los promedios de los tratamientos en orden ascendente, se tiene

$$\begin{aligned} \bar{y}_{1.} &= 9.8 \\ \bar{y}_{5.} &= 10.8 \\ \bar{y}_{2.} &= 15.4 \\ \bar{y}_{3.} &= 17.6 \\ \bar{y}_{4.} &= 21.6 \end{aligned}$$

El error estándar de cada promedio es $S_{\bar{y}_{i.}} = \sqrt{8.06/5} = 1.27$. En el conjunto de rangos significativos de la tabla VII del apéndice para 20 grados de libertad y $\alpha = 0.05$, se obtiene $r_{0.05}(2, 20) = 2.95$, $r_{0.05}(3, 20) = 3.10$, $r_{0.05}(4, 20) = 3.18$ y $r_{0.05}(5, 20) = 3.25$. Por lo tanto, los rangos de significación mínima son

$$\begin{aligned} R_2 &= r_{0.05}(2,\ 20)S_{\bar{y}_{i.}} = (2.95)(1.27) = 3.75 \\ R_3 &= r_{0.05}(3,\ 20)S_{\bar{y}_{i.}} = (3.10)(1.27) = 3.94 \\ R_4 &= r_{0.05}(4,\ 20)S_{\bar{y}_{i.}} = (3.18)(1.27) = 4.04 \\ R_5 &= r_{0.05}(5,\ 20)S_{\bar{y}_{i.}} = (3.25)(1.27) = 4.13 \end{aligned}$$

Los resultados de las comparaciones serían

$$\begin{aligned}
&4 \text{ vs. } 1\text{: } 21.6 - 9.8 = 11.8 > 4.13\ (R_5) \\
&4 \text{ vs. } 5\text{: } 21.6 - 10.8 = 10.8 > 4.04\ (R_4) \\
&4 \text{ vs. } 2\text{: } 21.6 - 15.4 = 6.2 > 3.94\ (R_3) \\
&4 \text{ vs. } 3\text{: } 21.6 - 17.6 = 4.0 > 3.75\ (R_2) \\
&3 \text{ vs. } 1\text{: } 17.6 - 9.8 = 7.8 > 4.04\ (R_4) \\
&3 \text{ vs. } 5\text{: } 17.6 - 10.8 = 6.8 > 3.95\ (R_3) \\
&3 \text{ vs. } 2\text{: } 17.6 - 15.4 = 2.2 < 3.75\ (R_2) \\
&2 \text{ vs. } 1\text{: } 15.4 - 9.8 = 5.6 > 3.94\ (R_3) \\
&2 \text{ vs. } 5\text{: } 15.4 - 10.8 = 4.6 > 3.75\ (R_2) \\
&5 \text{ vs. } 1\text{: } 10.8 - 9.8 = 1.0 < 3.75\ (R_2)
\end{aligned}$$

Por el análisis se observa que hay diferencias significativas entre todos los pares de medias con excepción de la 3 y la 2 y la 5 y la 1. En la figura 3-14 se muestra una gráfica en la que esas medias que no son significativamente diferentes aparecen subrayadas. Observe que en este ejemplo la prueba del rango múltiple de Duncan y el método LSD llevan a conclusiones idénticas.

....

$\bar{y}_{1.}$	$\bar{y}_{5.}$	$\bar{y}_{2.}$	$\bar{y}_{3.}$	$\bar{y}_{4.}$
9.8	10.8	15.4	17.6	21.6

Figura 3-14 Resultados de la prueba de rango múltiple de Duncan.

En la prueba del rango múltiple de Duncan se requiere una diferencia observada más grande para detectar pares significativamente diferentes de medias, cuando el número de medias incluidas en el grupo aumenta. De esta forma, en el ejemplo anterior $R_2 = 3.75$ (dos medias) mientras que $R_3 = 3.94$ (tres medias). Para dos medias, el valor crítico R_2 será exactamente igual al valor LSD de la prueba t. Los valores $r_\alpha(p, f)$ de la tabla VII del apéndice se eligen de tal modo que se obtenga un nivel de protección especificado. Es decir, cuando se comparan dos medias que están p pasos aparte, el **nivel de protección** es $(1-\alpha)^{p-1}$, donde α es el nivel de significación especificado para dos medias *adyacentes*. Por lo tanto, el índice de error de reportar al menos una diferencia significativa incorrecta entre dos medias es $1-(1-\alpha)^{p-1}$, cuando el tamaño del grupo es p. Por ejemplo, si $\alpha = 0.05$, entonces $1-(1-0.05)^1 = 0.05$ es el nivel de significación para comparar el par de medias adyacentes, $1-(1-0.05)^2 = 0.10$ es el nivel de significación para medias que están un paso aparte, y así sucesivamente.

En general, si el nivel de protección es α, las pruebas de las medias tienen un nivel de significación que es mayor o igual que α. Por consiguiente, el procedimiento de Duncan tiene una gran potencia; es decir, es muy eficaz para detectar diferencias entre medias cuando existen diferencias reales. Por esta razón, la prueba del rango múltiple de Duncan es muy popular.

La prueba de Newman-Keuls

Esta prueba fue creada por Newman [90]. Debido a que un interés renovado en la prueba de Newman fue generado por Keuls [64], al procedimiento se le llama la prueba de Newman-Keuls. Operacionalmente, el procedimiento es similar a la prueba del rango múltiple de Duncan, salvo porque las diferencias críticas entre las medias se calculan en una forma un tanto diferente. Específicamente, se calcula una serie de valores críticos

$$K_p = q_\alpha(p, f) S_{\bar{y}_{i.}} \qquad p = 2, 3, \ldots, a \tag{3-45}$$

donde $q_\alpha(p, f)$ es el punto porcentual α superior del rango studentizado para grupos de medias de tamaño p y con f grados de libertad del error. Una vez que se calculan los valores K_p con la ecuación 3-45, los pares de medias extremos en los grupos de tamaño p se comparan con K_p exactamente igual que en la prueba del rango múltiple de Duncan.

¿Qué método de comparación por pares debe usarse?

Ciertamente, una pregunta lógica en este punto es qué método de comparación por pares debe usarse. Desafortunadamente, no hay una respuesta precisa para esta pregunta, y los especialistas en estadística están con frecuencia en desacuerdo en cuanto a la utilidad de los diferentes procedimientos. Carmer y Swanson [24] han realizado estudios de simulación Montecarlo con varios procedimientos de comparaciones múltiples, incluyendo algunos que no se han considerado aquí. Estos autores reportan que el método de la diferencia significativa mínima es una prueba muy eficaz para detectar diferencias reales en las medias si se aplica *sólo después* de que la prueba F en el análisis de varianza sea significativa en 5%. Reportan asimismo un buen desempeño en la detección de diferencias reales con la prueba del rango múltiple de Duncan. Esto no es motivo de sorpresa, ya que estos dos métodos son los más poderosos de los que se han comentado aquí. Sin embargo, estos métodos no incluyen el índice de error en el modo del experimento. Debido a que el método de Tukey efectúa un control sobre el índice de error global, muchos experimentadores prefieren su uso.

La prueba de Newman-Keuls es más conservadora que la prueba del rango múltiple de Duncan por cuanto a que el índice de error tipo I es menor. Específicamente, el error tipo I del experimento es α para todas las pruebas que incluyen el mismo número de medias. Por consiguiente, debido a que α es por lo general bajo, la potencia de la prueba de Newman-Keuls casi siempre es menor que la de la prueba del rango múltiple de Duncan. Para demostrar que el procedimiento de Newman-Keuls lleva a una prueba con menor potencia que la prueba del rango múltiple de Duncan, se observa por una comparación de las ta-

blas VII y VIII del apéndice que para $p > 2$ se tiene $q_\alpha(p,f) > r_\alpha(p,f)$. Es decir, es "más difícil" declarar que un par de medias es significativamente diferente al utilizar la prueba de Newman-Keuls que cuando se usa el procedimiento de Duncan. Esto se ilustra a continuación para el caso en que $\alpha = 0.01, a = 8$ y $f = 20$:

p	2	3	4	5	6	7	8
$r_{0.01}(p, 20)$	4.02	4.22	4.33	4.40	4.47	4.53	4.58
$q_{0.01}(p, 20)$	4.02	4.64	5.02	5.29	5.51	5.69	5.84

Como se señaló antes, existen otros procedimientos de comparaciones múltiples. Algunos artículos que describen estos métodos son los de Miller [78], O'Neill y Wetherill [91] y Nelson [89]. También se recomienda el libro de Miller [77].

3-5.8 Comparación de medias de tratamientos con un control

En muchos experimentos, uno de los tratamientos es un **control,** y el analista se interesa en comparar cada una de las medias de los $a - 1$ tratamientos restantes con el control. Por lo tanto, sólo es necesario hacer $a - 1$ comparaciones. Un procedimiento para hacer estas comparaciones ha sido desarrollado por Dunnett [42]. Suponga que el tratamiento a es el control y que quieren probarse las hipótesis

$$H_0: \mu_i = \mu_a$$
$$H_1: \mu_i \neq \mu_a$$

para $i = 1, 2, ..., a - 1$. El procedimiento de Dunnett es una modificación de la prueba t común. Para cada hipótesis se calculan las diferencias observadas en las medias muestrales

$$|\bar{y}_{i.} - \bar{y}_{a.}| \qquad i = 1, 2, \ldots, a-1$$

La hipótesis nula H_0:$\mu_i = \mu_a$ se rechaza utilizando un índice α de error tipo I si

$$|\bar{y}_{i.} - \bar{y}_{a.}| > d_\alpha(a-1, f)\sqrt{MS_E\left(\frac{1}{n_i} + \frac{1}{n_a}\right)} \tag{3-46}$$

donde la constante $d_\alpha(a - 1, f)$ se da en la tabla IX del apéndice. (Pueden hacerse pruebas tanto de una como de dos colas.) Observe que α es el nivel de significación conjunto asociado con las $a - 1$ pruebas.

EJEMPLO 3-10 ....

Para ilustrar la prueba de Dunnett, considere el experimento del ejemplo 3-1, asumiendo que el tratamiento 5 es el control. En este ejemplo, $a = 5, a - 1 = 4, f = 20$ y $n_i = n = 5$. Con el nivel de 5%, en la tabla IX del apéndice se encuentra que $d_{0.05}(4, 20) = 2.65$. Por lo tanto, la diferencia crítica es

$$d_{0.05}(4, 20)\sqrt{\frac{2MS_E}{n}} = 2.65\sqrt{\frac{2(8.06)}{5}} = 4.76$$

(Observe que se trata de una simplificación de la ecuación 3-46 que resulta de un diseño balanceado.) Por lo tanto, cualquier media de los tratamientos que difiera del control por más de 4.76 se declararía significativamente diferente. Las diferencias observadas son

$$\begin{aligned}
1 \text{ vs. } 5: \bar{y}_{1.} - \bar{y}_{5.} &= 9.8 - 10.8 = -1.0 \\
2 \text{ vs. } 5: \bar{y}_{2.} - \bar{y}_{5.} &= 15.4 - 10.8 = 4.6 \\
3 \text{ vs. } 5: \bar{y}_{3.} - \bar{y}_{5.} &= 17.6 - 10.8 = 6.8 \\
4 \text{ vs. } 5: \bar{y}_{4.} - \bar{y}_{5.} &= 21.6 - 10.8 = 10.8
\end{aligned}$$

Sólo las diferencias $\bar{y}_{3.} - \bar{y}_{5.}$ y $\bar{y}_{4.} - \bar{y}_{5.}$ indican alguna diferencia significativa cuando se comparan con el control; por lo tanto, se concluye que $\mu_3 \neq \mu_5$ y $\mu_4 \neq \mu_5$.

Cuando se hace la comparación de los tratamientos con un control, una buena idea es usar más observaciones para el tratamiento de control (por ejemplo, n_a) que para los demás tratamientos (por ejemplo, n), suponiendo un número igual de observaciones para los $a - 1$ tratamientos restantes. El cociente n_a/n deberá elegirse de tal modo que sea aproximadamente igual a la raíz cuadrada del número total de tratamientos. Es decir, se elige $n_a/n = \sqrt{a}$.

3-6 MUESTRA DE SALIDA DE COMPUTADORA

Hay una gran cantidad de programas de computadora para apoyar el diseño experimental y la realización de análisis de varianza. En la figura 3-15 se muestra la salida de uno de estos programas, *Design-Expert*, utilizando los datos del experimento con un solo factor del ejemplo 3-1. La suma de cuadrados correspondiente al "Modelo" ("Model") es la $SS_{\text{Tratamientos}}$ usual de un diseño con un solo factor. Esa fuente se identifica adicionalmente como "A". Cuando hay más de un factor en el experimento, la suma de cuadrados ("Sum of Squares") del modelo se descompondrá en varias fuentes (A, B, etc.). Observe que el resumen del análisis de varianza de la parte superior de la salida de computadora contiene las sumas de cuadrados, los grados de libertad ("DF", degrees of freedom), los cuadrados medios ("Mean Square") y el estadístico de prueba F_0 ("F Value") acostumbrados. La columna "Prob > F" es el valor P (de hecho, el límite superior del valor P, ya que a las probabilidades menores que 0.0001 se les asigna el valor por omisión 0.0001).

Además del análisis de varianza básico, el programa presenta información adicional útil. La cantidad "R cuadrada" ("R-Squared") se define como

$$R^2 = \frac{SS_{\text{Modelo}}}{SS_{\text{Total}}} = \frac{475.76}{636.96} = 0.746923$$

y se interpreta en términos generales como la proporción de la variabilidad en los datos "explicada" por el modelo del análisis de varianza. Por lo tanto, en los datos para probar la resistencia de la fibra sintética, el factor "peso porcentual del algodón" explica cerca de 74.69% de la variabilidad en la resistencia a la tensión. Evidentemente, debe tenerse $0 \leq R^2 \leq 1$, siendo más deseables los valores más grandes. En la salida se presentan también otros estadísticos en R^2. R^2 "ajustada" ("Adj R-Squared") es una variante del estadístico R^2 común que refleja el número de factores presentes en el modelo. Puede ser un estadístico útil en experimentos más complejos en los que intervienen varios factores en el diseño, cuando quiere evaluarse el impacto de aumentar o disminuir el número de términos del modelo. "Desviación estándar" ("Std. Dev.") es la raíz cuadrada del cuadrado medio del error, $\sqrt{8.060} = 2.839$, y "C.V." es el coeficiente de variación, definido como $(\sqrt{MS_E} / \bar{y})100$. El coeficiente de variación mide la variabilidad no explicada o residual de los datos como un porcentaje de la media ("Mean") de la variable de respuesta. "PRESS" son las siglas de *Prediction Error Sum of Squares* (suma de cuadrados del error de predicción) y es una medida de la adecuación con que es posible que el modelo del experimento predecirá las respuestas en un *nuevo experimento*. Son deseables valores pequeños de PRESS. Alternativamente, puede calcularse una R^2 para predicciones con base en PRESS (más adelante se indicará cómo hacer esto). Esta R^2_{Pred} ("Pred R-Squared") para el problema tratado aquí es 0.6046, el cual no es irrazonable, considerando que el modelo explica cerca de 75% de la variabilidad del experimento en curso. El estadístico "Predicción adecuada" ("Adeq Precision") se calcula dividiendo la diferencia entre la respuesta predicha máxima y la respuesta predicha mínima por la desviación estándar promedio de todas las respuestas predichas. Son deseables valores grandes de esta cantidad, y los valores que exceden cuatro indican por lo general que el modelo tendrá un desempeño razonable en la predicción.

Utilice el mouse para posicionarse en una celda y su definición.

Response: Strength in psi
ANOVA for Selected Factorial Model
Analysis of variance table [Partial sum of squares]

Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	475.76	4	118.94	14.76	<0.0001	significativo
A	*475.76*	*4*	*118.94*	*14.76*	*<0.0001*	
Residual	161.20	20	8.06			
Lack of Fit	*0.000*	*0*				
Pure Error	*161.20*	*20*	*8.06*			
Cor Total	636.96	24				

El valor F del Modelo de 14.76 implica que el modelo es significativo. Sólo hay una probabilidad de 0.01% de que un "Valor F del Modelo" de esta magnitud pudiera ocurrir debido a ruido.

Los valores de "Prob > F" menores que 0.0500 indican que los términos del modelo son significativos.

En este caso A son términos significativos del modelo.
Los valores mayores que 0.1000 indican que los términos del modelo no son significativos.
Si hay muchos términos del modelo no significativos (sin contar los que se necesitan para apoyar la jerarquización), la reducción del modelo puede mejorarlo.

Std. Dev.	2.84	R-Squared	0.7469
Mean	15.04	Adj R-Squared	0.6963
C.V.	18.88	Pred R-Squared	0.6046
PRESS	251.88	Adeq Precision	9.294

La "R cuadrada predicha" de 0.6046 concuerda razonablemente con la "R cuadrada ajustada" de 0.6963. Una diferencia mayor que 0.20 entre la "R cuadrada predicha" y la "R cuadrada ajustada" indica un posible problema con el modelo y/o los datos.

"Precisión adecuada" mide la relación de la señal a ruido. Es deseable una relación mayor que 4. La relación de 9.294 indica una señal adecuada para usar este modelo para navegar el espacio del diseño.

Treatment Means (Adjusted, If Necessary)

	Estimated Mean	Standard Error
1-15	9.80	1.27
2-20	15.40	1.27
3-25	17.60	1.27
4-30	21.60	1.27
5-35	10.80	1.27

Treatment	Mean Difference	DF	Standard Error	t for H_0 Coeff=0	Prob > |t|
1 vs 2	−5.60	1	1.80	−3.12	0.0054
1 vs 3	−7.80	1	1.80	−4.34	0.0003
1 vs 4	−11.80	1	1.80	−6.57	<0.0001
1 vs 5	−1.00	1	1.80	−0.56	0.5838
2 vs 3	−2.20	1	1.80	−1.23	0.2347
2 vs 4	−6.20	1	1.80	−3.45	0.0025
2 vs 5	4.60	1	1.80	2.56	0.0186
3 vs 4	−4.00	1	1.80	−2.23	0.0375
3 vs 5	6.80	1	1.80	3.79	0.0012
4 vs 5	10.80	1	1.80	6.01	<0.0001

Los valores de "Prob > |t|" menores que 0.0500 indican que la diferencia en las medias de los dos tratamientos es significativa.
Los valores de "Prob > |t|" mayores que 0.1000 indican que la diferencia en las medias de los dos tratamientos no es significativa.

Figura 3-15 Salida de computadora de *Design-Expert* para el ejemplo 3-1.

Diagnostic Case Statistics

Standard Order	Actual Value	Predicted Value	Residual	Leverage	Student Residual	Cook's Distance	Outlier t
1	7.00	9.80	−2.80	0.200	−1.103	0.061	−1.109
2	7.00	9.80	−2.80	0.200	−1.103	0.061	−1.109
3	15.00	9.80	5.20	0.200	2.048	0.210	2.245
4	11.00	9.80	1.20	0.200	0.473	0.011	0.463
5	9.00	9.80	−0.80	0.200	−0.315	0.005	−0.308
6	12.00	15.40	−3.40	0.200	−1.339	0.090	−1.368
7	17.00	15.40	1.60	0.200	0.630	0.020	0.620
8	12.00	15.40	−3.40	0.200	−1.339	0.090	−1.368
9	18.00	15.40	2.60	0.200	1.024	0.052	1.025
10	18.00	15.40	2.60	0.200	1.024	0.052	1.025
11	14.00	17.60	−3.60	0.200	−1.418	0.100	−1.457
12	18.00	17.60	0.40	0.200	0.158	0.001	0.154
13	18.00	17.60	0.40	0.200	0.158	0.001	0.154
14	19.00	17.60	1.40	0.200	0.551	0.015	0.542
15	19.00	17.60	1.40	0.200	0.551	0.015	0.542
16	19.00	21.60	−2.60	0.200	−1.024	0.052	−1.025
17	25.00	21.60	3.40	0.200	1.339	0.090	1.368
18	22.00	21.60	0.40	0.200	0.158	0.001	0.154
19	19.00	21.60	−2.60	0.200	−1.024	0.052	−1.025
20	23.00	21.60	1.40	0.200	0.551	0.015	0.542
21	7.00	10.80	−3.80	0.200	−1.496	0.112	−1.548
22	10.00	10.80	−0.80	0.200	−0.315	0.005	−0.308
23	11.00	10.80	0.20	0.200	0.079	0.000	0.077
24	15.00	10.80	4.20	0.200	1.654	0.137	1.735
25	11.00	10.80	0.20	0.200	0.079	0.000	0.077

Proceder con las gráficas de diagnóstico (el icono siguiente en progresión). Asegurarse de examinar:

1) La gráfica de probabilidad normal de los residuales studentizados para verificar la normalidad de los residuales.
2) Los residuales studentizados contra los valores predichos para verificar la constante del error.
3) Los puntos atípicos t contra el orden de las corridas para buscar puntos atípicos, es decir, valores influyentes o importantes
4) La gráfica de Box-Cox para las transformaciones de potencia.

Si todos los estadísticos del modelo y las gráficas de diagnóstico están correctos, finalizar con el icono Model Graphs (Gráficas del Modelo).

Figura 3-15 (*Continuación.*)

Se hace la estimación de las medias ("Estimated Mean") de los tratamientos y se muestra el error estándar ("Standard Error") (o desviación estándar muestral de la media de cada tratamiento, $\sqrt{MS_E/n}$). Las diferencias entre pares de medias ("Mean Difference") de los tratamientos se investigan utilizando el método LSD de Fisher descrito en la sección 3-5.7.

El programa de computadora también calcula y despliega los residuales, según se definen en la ecuación 3-16. El programa producirá también todas las gráficas de los residuales que se comentaron en la sección 3-4. En la salida se muestran asimismo varios diagnósticos residuales más. Algunos de ellos se revisarán más adelante.

Por último, observe que el programa de computadora incluye también algunas guías para hacer la interpretación. Esta información "aconsejable" es muy común en muchos paquetes de estadística para computadoras personales. Al leer estas guías, recuerde que están escritas en términos muy generales, y quizá no se ajusten exactamente a los requerimientos de redacción del reporte de un experimentador particular. Esta salida aconsejable puede ser eliminada por el usuario.

3-7 DETERMINACIÓN DEL TAMAÑO DE LA MUESTRA

En cualquier problema de diseño experimental, una decisión crítica es la elección del tamaño de la muestra; es decir, determinar el número de réplicas que deben correrse. En general, si el experimentador tiene interés en detectar efectos pequeños, se necesitan más réplicas que cuando el experimentador se interesa en detectar efectos grandes. En esta sección se analizan varios enfoques para determinar el tamaño de la muestra. Aun cuando la revisión se centra en un diseño con un solo factor, la mayoría de los métodos pueden usarse en situaciones experimentales más complejas.

3-7.1 Curvas de operación característica

Recuerde que una **curva de operación característica** es una gráfica de la probabilidad del error tipo II de una prueba estadística para un tamaño de la muestra particular contra un parámetro que refleja la medida en que la hipótesis nula es falsa. El experimentador puede usar estas curvas como guía en la selección del número de réplicas para que el diseño sea sensible a diferencias potenciales importantes en los tratamientos.

Se considera la probabilidad del error tipo II del modelo con efectos fijos para el caso en que se usa el mismo tamaño de las muestras en cada tratamiento, por ejemplo

$$\begin{aligned}\beta &= 1 - P\{\text{Rechazar } H_0 \mid H_0 \text{ es falsa}\} \\ &= 1 - P\{F_0 > F_{\alpha,a-1,N-a} \mid H_0 \text{ es falsa}\}\end{aligned} \qquad (3\text{-}47)$$

Para evaluar el enunciado de probabilidad de la ecuación 3-47, es necesario conocer cuál es la distribución del estadístico de prueba F_0 si la hipótesis nula es falsa. Puede demostrarse que, si H_0 es falsa, el estadístico $F_0 = MS_{\text{Tratamientos}}/MS_E$ se distribuye como una variable aleatoria ***F* no central** con $a - 1$ y $N - a$ grados de libertad y parámetro de no centralidad δ. Si $\delta = 0$, la distribución F no central se convierte en la distribución F (central) común.

Las curvas de operación característica que se presentan en la parte V del apéndice se usan para evaluar el enunciado de probabilidad de la ecuación 3-47. En estas curvas se grafica la probabilidad del error tipo II (β) contra un parámetro Φ, donde

$$\Phi^2 = \frac{n\sum_{i=1}^{a} \tau_i^2}{a\sigma^2} \qquad (3\text{-}48)$$

La cantidad Φ^2 está relacionada con el parámetro de no centralidad δ. Se cuenta con curvas para $\alpha = 0.05$ y $\alpha = 0.01$ y un rango de grados de libertad para el numerador y el denominador.

Al usar las curvas de operación característica, el experimentador debe especificar el parámetro Φ. Con frecuencia es difícil hacer esto en la práctica. Una manera de determinar Φ es elegir los valores reales de las medias de los tratamientos para los que querría rechazarse la hipótesis nula con una alta probabilidad. Por lo tanto, si $\mu_1, \mu_2, ..., \mu_a$ son las medias de los tratamientos especificadas, la τ_i de la ecuación 3-48 se encuentra como $\tau_i = \mu_i - \bar{\mu}$, donde $\bar{\mu} = (1/a)\Sigma_{i=1}^{a}\mu_i$ es el promedio de las medias de los tratamientos individuales. Se requiere asimismo una estimación de σ^2. En ocasiones se cuenta con este valor por experiencia previa, un experimento anterior o una prueba preliminar (como se sugirió en el capítulo 1), o por una estimación discrecional. Cuando no se tiene la seguridad acerca del valor de σ^2, los tamaños de las muestras podrían determinarse para un rango de valores posibles de σ^2, a fin de estudiar el efecto de este parámetro sobre el tamaño de la muestra requerido, antes de hacer la elección final.

EJEMPLO 3-11 ....

Considere el experimento de la resistencia a la tensión descrito en el ejemplo 3-1. Suponga que el experimentador está interesado en rechazar la hipótesis nula con una probabilidad de al menos 0.90 si las medias de los cinco tratamientos son

$$\mu_1=11 \qquad \mu_2=12 \qquad \mu_3=15 \qquad \mu_4=18 \qquad \text{y} \qquad \mu_5=19$$

Planea utilizar $\alpha = 0.01$. En este caso, puesto que $\Sigma_{i=1}^{5}\mu_i = 75$, se tiene $\bar{\mu} = (1/5)75 = 15$ y

$$\begin{aligned}
\tau_1 &= \mu_1 - \bar{\mu} = 11-15= -4\\
\tau_2 &= \mu_2 - \bar{\mu} = 12-15= -3\\
\tau_3 &= \mu_3 - \bar{\mu} = 15-15= \ \ 0\\
\tau_4 &= \mu_4 - \bar{\mu} = 18-15= \ \ 3\\
\tau_5 &= \mu_5 - \bar{\mu} = 19-15= \ \ 4
\end{aligned}$$

Por lo tanto, $\Sigma_{i=1}^{5}\tau_i^2 = 50$. Suponga que el experimentador piensa que la desviación estándar de la resistencia a la tensión con cualquier nivel particular del peso porcentual del algodón no será mayor que $\sigma=3$ psi. Entonces, al utilizar la ecuación 3-48, se tiene

$$\Phi^2 = \frac{n\sum_{i=1}^{5}\tau_i^2}{a\sigma^2} = \frac{n(50)}{5(3)^2} = 1.11n$$

Se usa la curva de operación característica para $a-1 = 5-1 = 4$ con $N-a = a(n-1) = 5(n-1)$ grados de libertad del error y $\alpha = 0.01$ (ver la parte V del apéndice). Como primera conjetura para el tamaño de la muestra requerido, se prueba con $n = 4$ réplicas. Esto produce $\Phi^2 = 1.11(4) = 4.44$, $\Phi = 2.11$ y $5(3) = 15$ grados de libertad del error. Por consiguiente, en la parte V se encuentra que $\beta \simeq 0.30$. Por lo tanto, la potencia de la prueba es aproximadamente $1-\beta = 1-0.30 = 0.70$, que es menor que el 0.90 requerido, por lo que se concluye que $n = 4$ réplicas no son suficientes. Procediendo de manera similar, puede construirse la siguiente tabla:

n	Φ^2	Φ	$a(n-1)$	β	Potencia $(1-\beta)$
4	4.44	2.11	15	0.30	0.70
5	5.55	2.36	20	0.15	0.85
6	6.66	2.58	25	0.04	0.96

Por lo tanto, deben realizarse al menos $n = 6$ réplicas para obtener una prueba con la potencia requerida.

....

El único problema con este enfoque para usar las curvas de operación característica es que por lo general es difícil seleccionar el conjunto de las medias de los tratamientos en el que se basará la decisión del tamaño de la muestra. Un enfoque alternativo es seleccionar un tamaño de la muestra tal que si la diferencia entre las medias de dos tratamientos cualesquiera excede un valor especificado, la hipótesis nula deberá rechazarse. Si la diferencia entre las medias de dos tratamientos cualesquiera es tan grande como D, puede demostrarse que el valor mínimo de Φ^2 es

$$\Phi^2 = \frac{nD^2}{2a\sigma^2} \tag{3-49}$$

Puesto que éste es un valor mínimo de Φ^2, el tamaño de la muestra correspondiente que se obtiene de la curva de operación característica es un valor conservador; es decir, proporciona una potencia al menos tan grande como la que especificó el experimentador.

Para ilustrar este enfoque, suponga que en el experimento de la resistencia a la tensión del ejemplo 3-1, el experimentador quisiera rechazar la hipótesis nula con una probabilidad de al menos 0.90 si las medias de dos tratamientos cualesquiera difieren hasta en 10 psi. Entonces, suponiendo que $\sigma = 3$ psi, se encuentra que el valor mínimo de Φ^2 es

$$\Phi^2 = \frac{n(10)^2}{2(5)(3^2)} = 1.11n$$

y, por el análisis del ejemplo 3-11, se concluye que se necesitan $n = 6$ réplicas para obtener la sensibilidad deseada cuando $\alpha = 0.01$.

3-7.2 Especificación de un incremento de la desviación estándar

Este enfoque es útil en ocasiones para elegir el tamaño de la muestra. Si las medias de los tratamientos no difieren, la desviación estándar de una observación elegida al azar es σ. Sin embargo, si las medias de los tratamientos son diferentes, la desviación estándar de una observación elegida al azar es

$$\sqrt{\sigma^2 + \left(\sum_{i=1}^{a} \tau_i^2 / a\right)}$$

Si se escoge un porcentaje P para el incremento de la desviación estándar de una observación, más allá del cual quiera rechazarse la hipótesis de que las medias de todos los tratamientos son iguales, esto es equivalente a escoger

$$\frac{\sqrt{\sigma^2 + \left(\sum_{i=1}^{a} \tau_i^2 / a\right)}}{\sigma} = 1 + 0.01P \quad (P = \text{por ciento})$$

o

$$\frac{\sqrt{\left(\sum_{i=1}^{a} \tau_i^2 / a\right)}}{\sigma} = \sqrt{(1+0.01P)^2 - 1}$$

de donde

$$\Phi = \frac{\sqrt{\sum_{i=1}^{a} \tau_i^2 / a}}{\sigma / \sqrt{n}} = \sqrt{(1+0.01P)^2 - 1}(\sqrt{n}) \tag{3-50}$$

Por lo tanto, para un valor especificado de P, Φ puede calcularse con la ecuación 3-50 y después usar las curvas de operación característica de la parte V del apéndice para determinar el tamaño de la muestra requerido.

Por ejemplo, en el experimento de la resistencia a la tensión del ejemplo 3-1, suponga que se desea detectar un incremento de la desviación estándar de 20% con una probabilidad de al menos 0.90 y α = 0.05. Entonces

$$\Phi = \sqrt{(1.2)^2 - 1}(\sqrt{n}) = 0.66\sqrt{n}$$

La referencia a las curvas de operación característica indica que se necesita $n = 9$ para obtener la sensibilidad deseada.

3-7.3 Método para estimar el intervalo de confianza

En este enfoque se supone que el experimentador quiere expresar los resultados finales en términos de intervalos de confianza y que está dispuesto a especificar por anticipado cuál es el ancho que desea para estos intervalos de confianza. Por ejemplo, suponga que en el experimento de la resistencia a la tensión del ejemplo 3-1 se quiere que un intervalo de confianza de 95% para la diferencia en la resistencia a la tensión media de dos pesos porcentuales del algodón cualesquiera sea ±5 psi y una estimación previa de σ es 3. Entonces, al utilizar la ecuación 3-13, se encuentra que la precisión del intervalo de confianza es

$$\pm t_{\alpha/2, N-a}\sqrt{\frac{2MS_E}{n}}$$

Suponga que se prueba con $n = 5$ réplicas. Entonces, al usar $\sigma^2 = 3^2 = 9$ como una estimación de MS_E, la precisión del intervalo de confianza es

$$\pm 2.086\sqrt{\frac{2(9)}{5}} = \pm 3.96$$

que es más preciso que el requerimiento. Al probar con $n = 4$ se obtiene

$$\pm 2.132\sqrt{\frac{2(9)}{4}} = \pm 4.52$$

Al probar con $n = 3$ se obtiene

$$\pm 2.228\sqrt{\frac{2(9)}{3}} = \pm 5.46$$

Evidentemente, $n = 4$ es el tamaño de la muestra menor que llevará a la precisión deseada.

El nivel de significación consignado en el ejemplo anterior se aplica a un solo intervalo de confianza. Sin embargo, puede usarse el mismo enfoque general si el experimentador desea especificar de antemano un *conjunto* de intervalos de confianza acerca del cual se hace un **enunciado de confianza simultáneo** o **conjunto** (ver los comentarios acerca de los intervalos de confianza simultáneos de la sección 3-3.3). Además, los intervalos de confianza podrían construirse con respecto a contrastes más generales en las medias de los tratamientos, que la comparación por pares ilustrada antes.

3-8 IDENTIFICACIÓN DE EFECTOS DE DISPERSIÓN

Nos hemos enfocado aquí en el uso del análisis de varianza y de otros métodos relacionados para determinar los niveles del factor que resultan en diferencias entre las medias de los tratamientos o los niveles del factor. Se acostumbra referirse a estos efectos como **efectos de localización.** Cuando ocurrió la desigual-

Tabla 3-12 Datos del experimento de fundición

Algoritmo para controlar la proporción	Observaciones					
	1	2	3	4	5	6
1	4.93(0.05)	4.86(0.04)	4.75(0.05)	4.95(0.06)	4.79(0.03)	4.88(0.05)
2	4.85(0.04)	4.91(0.02)	4.79(0.03)	4.85(0.05)	4.75(0.03)	4.85(0.02)
3	4.83(0.09)	4.88(0.13)	4.90(0.11)	4.75(0.15)	4.82(0.08)	4.90(0.12)
4	4.89(0.03)	4.77(0.04)	4.94(0.05)	4.86(0.05)	4.79(0.03)	4.76(0.02)

dad de la varianza con los diferentes niveles del factor, se utilizaron transformaciones para estabilizar la varianza y mejorar así las inferencias hechas sobre los efectos de localización. Sin embargo, en algunos problemas el interés se centra en descubrir si los diferentes niveles del factor afectan la **variabilidad;** es decir, el interés está en descubrir **efectos de dispersión** potenciales. Esto ocurrirá siempre que la desviación estándar, la varianza o cualquier otra medida de la variabilidad se use como variable de respuesta.

Para ilustrar estos conceptos, considere los datos de la tabla 3-12, los cuales se obtuvieron de un experimento diseñado en una fundición de aluminio. El aluminio se produce combinando alúmina con otros ingredientes en una celda de reacción y aplicando calor al hacer pasar una corriente eléctrica a través de la celda. La alúmina se agrega de manera continua a la celda para mantener la proporción apropiada de la misma con respecto a los otros ingredientes. En este experimento se investigaron cuatro algoritmos para controlar la proporción. Las variables de respuesta estudiadas se relacionaron con el voltaje de la celda. Específicamente, un sensor registra el voltaje de la celda varias veces cada segundo, produciendo miles de mediciones del voltaje durante cada corrida del experimento. Los ingenieros del proceso decidieron usar como variables de respuesta el voltaje promedio y la desviación estándar del voltaje de la celda (indicado entre paréntesis) en la corrida experimental. El voltaje promedio es importante porque afecta la temperatura de la celda, y la desviación estándar del voltaje (llamada "ruido del crisol" por los ingenieros del proceso) es importante porque afecta la eficiencia global de la celda.

Se llevó a cabo un análisis de varianza para determinar si los diferentes algoritmos para controlar la proporción afectan el voltaje promedio de la celda. Éste reveló que el algoritmo para controlar la proporción no tuvo ningún **efecto de localización;** es decir, al cambiar los algoritmos para controlar la proporción no hubo ningún cambio en el voltaje promedio de la celda. (Referirse al problema 3-28.)

Para investigar los efectos de dispersión, lo mejor suele ser utilizar

$$\log(s) \quad \text{o} \quad \log(s^2)$$

como variable de respuesta, ya que la transformación logarítmica es eficaz para estabilizar la variabilidad en la distribución de la desviación estándar muestral. Puesto que todas las desviaciones estándar del voltaje del crisol son menores que la unidad, se usará

$$y = -\ln(s)$$

como la variable de respuesta. En la tabla 3-13 se presenta el análisis de varianza para esta respuesta, el logaritmo natural del "ruido del crisol". Observe que la elección de un algoritmo para controlar la proporción afecta el ruido del crisol; es decir, el algoritmo para controlar la proporción tiene un **efecto de dis-**

Tabla 3-13 Análisis de varianza del logaritmo natural del ruido del crisol

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0	Valor P
Algoritmo para controlar la proporción	6.166	3	2.055	21.96	<0.001
Error	1.872	20	0.094		
Total	8.038	23			

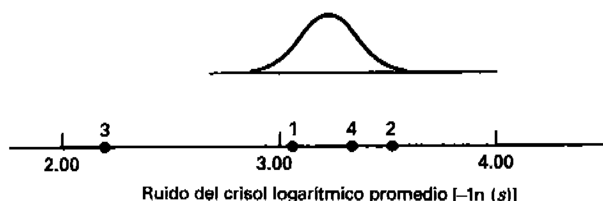


Figura 3-16 Ruido del crisol logarítmico promedio [$-\ln(s)$] de cuatro algoritmos para controlar la proporción en relación con una distribución t escalada con factor de escalamiento $\sqrt{MS_E / n} = \sqrt{0.094/6} = 0.125$.

persión. Las pruebas estándares de la adecuación del modelo, incluyendo las gráficas de probabilidad normal de los residuales, indican que no hay problemas con la validez del experimento. (Referirse al problema 3-29.)

En la figura 3-16 se grafica el logaritmo promedio del ruido del crisol de cada algoritmo para controlar la proporción y se presenta también una distribución t escalada que se usa como **distribución de referencia** para discriminar entre los algoritmos de la proporción. Esta gráfica revela con toda claridad que el algoritmo 3 para controlar la proporción produce más ruido del crisol o una desviación estándar del voltaje de la celda mayor que los otros algoritmos. No parece haber gran diferencia entre los algoritmos 1, 2 y 4.

3-9 EL ENFOQUE DE REGRESIÓN PARA EL ANÁLISIS DE VARIANZA

Se ha ofrecido un desarrollo intuitivo o heurístico del análisis de varianza. Sin embargo, es posible presentar un desarrollo más formal. El método será de utilidad más adelante para entender los fundamentos del análisis estadístico de diseños más complejos. Llamada la **prueba general de significación de la regresión,** el procedimiento consiste en esencia en encontrar la reducción en la suma de cuadrados total para ajustar el modelo con todos los parámetros incluidos y la reducción en la suma de cuadrados cuando el modelo se restringe a la hipótesis nula. La diferencia entre estas dos sumas de cuadrados es la suma de cuadrados de los tratamientos con la que puede realizarse la prueba de la hipótesis nula. El procedimiento requiere los estimadores de mínimos cuadrados de los parámetros en el modelo del análisis de varianza. Se dieron ya (en la sección 3-3.3) las estimaciones de estos parámetros; sin embargo, ahora se presenta un desarrollo formal.

3-9.1 Estimación de mínimos cuadrados de los parámetros del modelo

Se desarrollan ahora los estimadores de los parámetros en el modelo con un solo factor

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij}$$

utilizando el método de mínimos cuadrados. Para encontrar los estimadores de mínimos cuadrados de μ y τ_i, primero se forma la suma de cuadrados de los errores

$$L = \sum_{i=1}^{a}\sum_{j=1}^{n} \varepsilon_{ij}^2 = \sum_{i=1}^{a}\sum_{j=1}^{n} (y_{ij} - \mu - \tau_i)^2 \tag{3-51}$$

y se eligen después los valores de μ y τ_i, por ejemplo $\hat{\mu}$ y $\hat{\tau}_i$, que minimicen L. Los valores adecuados serían las soluciones de las $a + 1$ ecuaciones simultáneas

$$\left.\frac{\partial L}{\partial \mu}\right|_{\hat{\mu},\hat{\tau}_i} = 0$$

$$\left.\frac{\partial L}{\partial \tau_i}\right|_{\hat{\mu},\hat{\tau}_i} = 0 \qquad i = 1, 2, \ldots, a$$

Al derivar la ecuación 3-51 con respecto a μ y τ_i y al igualar con cero se obtiene

$$-2\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \hat{\mu} - \hat{\tau}_i) = 0$$

y

$$-2\sum_{j=1}^{n}(y_{ij} - \hat{\mu} - \hat{\tau}_i) = 0 \qquad i = 1, 2, \ldots, a$$

de la que, después de simplificar, se obtiene

$$\begin{aligned} N\hat{\mu} + n\hat{\tau}_1 + n\hat{\tau}_2 + \cdots + n\hat{\tau}_a &= y_{..} \\ n\hat{\mu} + n\hat{\tau}_1 \qquad\qquad &= y_{1.} \\ n\hat{\mu} \qquad + n\hat{\tau}_2 \qquad &= y_{2.} \\ \vdots \qquad\qquad &\quad \vdots \\ n\hat{\mu} \qquad\qquad + n\hat{\tau}_a &= y_{a.} \end{aligned} \tag{3-52}$$

A las $a + 1$ ecuaciones (ecuación 3-52) con $a + 1$ incógnitas se les llama las **ecuaciones normales de mínimos cuadrados.** Observe que si se suman la últimas a ecuaciones normales, se obtiene la primera ecuación normal. Por lo tanto, las ecuaciones normales no son linealmente independientes, y no existe una solución única para $\mu, \tau_1, \ldots, \tau_a$. Esta dificultad puede superarse mediante varios métodos. Puesto que los efectos de los tratamientos se han definido como desviaciones de la media global, parece razonable aplicar la **restricción**

$$\sum_{i=1}^{a} \hat{\tau}_i = 0 \tag{3-53}$$

Utilizando esta restricción, se obtiene como solución de las ecuaciones normales

$$\begin{aligned} \hat{\mu} &= \bar{y}_{..} \\ \hat{\tau}_i &= \bar{y}_{i.} - \bar{y}_{..} \qquad i = 1, 2, \ldots, a \end{aligned} \tag{3-54}$$

Evidentemente, esta solución no es única y depende de la restricción (ecuación 3-53) que se ha elegido. Al principio esto puede parecer desafortunado porque dos experimentadores diferentes podrían analizar los mismos datos y obtener resultados diferentes si aplican restricciones diferentes. Sin embargo, ciertas **funciones** del parámetro del modelo *son* estimadas de manera única, independientemente de la restricción. Algunos ejemplos son $\tau_i - \tau_j$, que se estimaría con $\hat{\tau}_i - \hat{\tau}_j = \bar{y}_{i.} - \bar{y}_{j.}$, y la media del tratamiento i-ésimo $\mu_i = \mu + \tau_i$, que se estimaría con $\hat{\mu}_i = \hat{\mu} + \hat{\tau}_i = \bar{y}_{i.}$.

Puesto que el interés se encuentra generalmente en las diferencias entre los efectos de los tratamientos y no en sus valores reales, no produce preocupación alguna que τ_i no pueda estimarse de manera úni-

ca. En general, cualquier función de los parámetros del modelo que sea una combinación lineal del miembro del lado izquierdo de las ecuaciones normales (ecuaciones 3-52) puede estimarse de manera única. A las funciones que se estiman de manera única independientemente de la restricción que se use se les llama **funciones estimables.** Para más información, ver el material suplementario del texto de este capítulo. Nos encontramos listos para usar estas estimaciones de los parámetros en un desarrollo general del análisis de varianza.

3-9.2 Prueba general de significación de la regresión

Una parte fundamental de este procedimiento es escribir las ecuaciones normales del modelo. Estas ecuaciones siempre podrán obtenerse formando la función de mínimos cuadrados y derivándola con respecto a cada parámetro desconocido, como se hizo en la sección 3-9.1. Sin embargo, se cuenta con un método más sencillo. Las reglas siguientes permiten escribir directamente las ecuaciones normales del modelo de *cualquier* diseño experimental:

REGLA 1. Hay una ecuación normal para cada parámetro del modelo que va a estimarse.

REGLA 2. El miembro derecho de cualquier ecuación normal es sólo la suma de todas las observaciones que contienen el parámetro asociado con esa ecuación normal particular.

Para ilustrar esta regla, considere el modelo con un solo factor. La primera ecuación normal corresponde al parámetro μ; por lo tanto, el miembro derecho es $y_{..}$, ya que *todas* las observaciones incluyen a μ.

REGLA 3. El miembro izquierdo de cualquier ecuación normal es la suma de todos los parámetros del modelo, donde cada parámetro está multiplicado por el número de veces que aparece en el total del miembro derecho. Los parámetros se escriben con un acento circunflejo (ˆ) para indicar que son *estimadores* y no los verdaderos valores de los parámetros.

Por ejemplo, considere la primera ecuación normal en un experimento con un solo factor. De acuerdo con las reglas anteriores, ésta sería

$$N\hat{\mu}+n\hat{\tau}_1+n\hat{\tau}_2+\cdots+n\hat{\tau}_a=y_{..}$$

porque μ aparece en las N observaciones, τ_1 sólo aparece en las n observaciones hechas bajo el primer tratamiento, τ_2 aparece sólo en las n observaciones tomadas bajo el segundo tratamiento, etc. Por la ecuación 3-52 se verifica que la ecuación presentada arriba es correcta. La segunda ecuación normal correspondería a τ_1 y es

$$n\hat{\mu}+n\hat{\tau}_1=y_{1.}$$

porque sólo las observaciones del primer tratamiento contienen a τ_1 (esto da $y_{1.}$ como miembro derecho), μ y τ_1 aparecen exactamente n veces en $y_{1.}$, y todas las demás τ_i aparecen cero veces. En general, el miembro izquierdo de cualquier ecuación normal es el valor esperado del miembro derecho.

Ahora bien, considere encontrar la reducción en la suma de cuadrados ajustando un modelo particular a los datos. Al ajustar un modelo a los datos se "explica" parte de la variabilidad; es decir, la variabilidad no explicada se reduce en cierta cantidad. La reducción en la variabilidad no explicada es siempre la suma de las estimaciones de los parámetros, cada una de ellas multiplicada por el segundo miembro de la

ecuación normal que corresponde al parámetro específico. Por ejemplo, en un experimento con un solo factor, la reducción debida al ajuste del modelo completo $y_{ij} = \mu + \tau_i + \varepsilon_{ij}$ es

$$\begin{aligned} R(\mu, \tau) &= \hat{\mu} y_{..} + \hat{\tau}_1 y_{1.} + \hat{\tau}_2 y_{2.} + \cdots + \hat{\tau}_a y_{a.} \\ &= \hat{\mu} y_{..} + \sum_{i=1}^{a} \hat{\tau}_i y_{i.} \end{aligned} \tag{3-55}$$

La notación $R(\mu, \tau)$ significa la reducción en la suma de cuadrados a partir del ajuste del modelo que contiene a μ y $\{\tau_i\}$. A $R(\mu, \tau)$ se le llama en ocasiones la suma de cuadrados "de regresión" del modelo completo $y_{ij} = \mu + \tau_i + \varepsilon_{ij}$. El número de grados de libertad asociado con una reducción en la suma de cuadrados, tal como $R(\mu, \tau)$, siempre es igual al número de ecuaciones normales linealmente independientes. El resto de la variabilidad no explicada por el modelo se encuentra con

$$SS_E = \sum_{i=1}^{a} \sum_{j=1}^{n} y_{ij}^2 - R(\mu, \tau) \tag{3-56}$$

Esta cantidad se usa en el denominador del estadístico de prueba de $H_0: \tau_1 = \tau_2 = \cdots = \tau_a = 0$.

A continuación se ilustra la prueba general de significación de la regresión para un experimento con un solo factor y se demuestra que produce el análisis de varianza de un solo factor común. El modelo es $y_{ij} = \mu + \tau_i + \varepsilon_{ij}$, y las ecuaciones normales se encuentran con las reglas anteriores como

$$\begin{aligned} N\hat{\mu} + n\hat{\tau}_1 + n\hat{\tau}_2 + \cdots + n\hat{\tau}_a &= y_{..} \\ n\hat{\mu} + n\hat{\tau}_1 \qquad\qquad\qquad &= y_{1.} \\ n\hat{\mu} \qquad + n\hat{\tau}_2 \qquad\qquad &= y_{2.} \\ \vdots \qquad\qquad &\ \ \vdots \\ n\hat{\mu} \qquad\qquad\qquad + n\hat{\tau}_a &= y_{a.} \end{aligned}$$

Compare estas ecuaciones normales con las que se obtuvieron en la ecuación 3-52.

Al aplicar la restricción $\Sigma_{i=1}^{a} \hat{\tau}_1 = 0$, los estimadores de μ y τ_i son

$$\hat{\mu} = \bar{y}_{..} \qquad \hat{\tau}_i = \bar{y}_{i.} - \bar{y}_{..} \qquad i = 1, 2, \ldots, a$$

La reducción en la suma de cuadrados debida al ajuste de este modelo completo se encuentra con la ecuación 3-55 como

$$\begin{aligned} R(\mu, \tau) &= \hat{\mu} y_{..} + \sum_{i=1}^{a} \hat{\tau}_i y_{i.} \\ &= (\bar{y}_{..}) y_{..} + \sum_{i=1}^{a} (\bar{y}_{i.} - \bar{y}_{..}) y_{i.} \\ &= \frac{y_{..}^2}{N} + \sum_{i=1}^{a} \bar{y}_{i.} y_{i.} - \bar{y}_{..} \sum_{i=1}^{a} y_{i.} \\ &= \sum_{i=1}^{a} \frac{y_{i.}^2}{n} \end{aligned}$$

que tiene a grados de libertad porque hay a ecuaciones normales linealmente independientes. La suma de cuadrados del error es, por la ecuación 3-56,

$$SS_E = \sum_{i=1}^{a}\sum_{j=1}^{n} y_{ij}^2 - R(\mu, \tau)$$

$$= \sum_{i=1}^{a}\sum_{j=1}^{n} y_{ij}^2 - \sum_{i=1}^{a} \frac{y_{i.}^2}{n}$$

y tiene $N - a$ grados de libertad.

Para encontrar la suma de cuadrados que resulta de los efectos de los tratamientos (el $\{\tau_i\}$), se considera que el modelo se restringe a la hipótesis nula; es decir, $\tau_i = 0$ para toda i. El modelo reducido es $y_{ij} = \mu + \varepsilon_{ij}$. Hay una sola ecuación normal para este modelo:

$$N\hat{\mu} = y_{..}$$

y el estimador de μ es $\hat{\mu} = \bar{y}_{..}$. Por lo tanto, la reducción en la suma de cuadrados que resulta de ajustar el modelo reducido que sólo contiene a μ es

$$R(\mu) = (\bar{y}_{..})(y_{..}) = \frac{y_{..}^2}{N}$$

Puesto que hay una sola ecuación normal para este modelo reducido, $R(\mu)$ tiene un grado de libertad. La suma de cuadrados debida al $\{\tau_i\}$, dado que μ ya está incluida en el modelo, es la diferencia entre $R(\mu, \tau)$ y $R(\mu)$, que es

$$R(\tau|\mu) = R(\mu, \tau) - R(\mu)$$

$$= R(\text{Modelo Completo}) - R(\text{Modelo Reducido})$$

$$= \frac{1}{n}\sum_{i=1}^{a} y_{i.}^2 - \frac{y_{..}^2}{N}$$

con $a - 1$ grados de libertad, que por la ecuación 3-9 se identifica como $SS_{\text{Tratamientos}}$. Estableciendo el supuesto de normalidad usual, el estadístico apropiado para probar H_0: $\tau_1 = \tau_2 = \cdots = \tau_a = 0$ es

$$F_0 = \frac{R(\tau|\mu)/(a-1)}{\left[\sum_{i=1}^{a}\sum_{j=1}^{n} y_{ij}^2 - R(\mu, \tau)\right]/(N-a)}$$

que se distribuye como $F_{a-1, N-a}$ bajo la hipótesis nula. Se trata, desde luego, del estadístico de prueba para el análisis de varianza de un solo factor.

3-10 MÉTODOS NO PARAMÉTRICOS EN EL ANÁLISIS DE VARIANZA

3-10.1 La prueba de Kruskal-Wallis

En situaciones en las que el supuesto de normalidad no está justificado, el experimentador quizá quiera usar un procedimiento alternativo del análisis de varianza con la prueba F que no dependa de este su-

puesto. Kruskal y Wallis [68] han desarrollado este procedimiento. La prueba de Kruskal-Wallis se usa para probar la hipótesis nula de que los a tratamientos son idénticos contra la hipótesis alternativa de que algunos de los tratamientos generan observaciones que son mayores que otras. Debido a que el procedimiento está diseñado para ser sensible al probar las diferencias en las medias, en ocasiones es conveniente considerar la prueba de Kruskal-Wallis como una prueba de la igualdad de las medias de los tratamientos. La prueba de Kruskal-Wallis es una **alternativa no paramétrica** del análisis de varianza usual.

Para realizar la prueba de Kruskal-Wallis, primero se hace la clasificación en rangos de las y_{ij} observaciones en orden ascendente y cada observación se reemplaza con su rango, por ejemplo R_{ij}, asignándole a la observación menor el rango 1. En el caso de empates (observaciones que tienen el mismo valor), se asigna el rango promedio a cada una de las observaciones empatadas. Sea $R_{i.}$ la suma de los rangos del tratamiento i-ésimo. El estadístico de prueba es

$$H = \frac{1}{S^2}\left[\sum_{i=1}^{a} \frac{R_{i.}^2}{n_i} - \frac{N(N+1)^2}{4}\right] \quad (3\text{-}57)$$

donde n_i es el número de observaciones del tratamiento i-ésimo, N es el número total de observaciones y

$$S^2 = \frac{1}{N-1}\left[\sum_{i=1}^{a}\sum_{j=1}^{n_i} R_{ij}^2 - \frac{N(N+1)^2}{4}\right] \quad (3\text{-}58)$$

Observe que S^2 es sólo la varianza de los rangos. Si no hay empates, $S^2 = N(N + 1)/12$, y el estadístico de prueba se simplifica a

$$H = \frac{12}{N(N+1)}\sum_{i=1}^{a} \frac{R_{i.}^2}{n_i} - 3(N+1) \quad (3\text{-}59)$$

Cuando el número de empates es moderado, habrá pequeñas diferencias entre las ecuaciones 3-57 y 3-59, y puede usarse la forma más simple (ecuación 3-59). Si las n_i son razonablemente grandes, por ejemplo $n_i \geq 5$, H se distribuye aproximadamente como χ^2_{a-1} bajo la hipótesis nula. Por lo tanto, si

$$H > \chi^2_{\alpha,a-1}$$

la hipótesis nula se rechaza. También podría usarse el enfoque del valor P.

EJEMPLO 3-12

En la tabla 3-14 se muestran los datos del ejemplo 3-1 y sus rangos correspondientes. Puesto que hay un número bastante grande de empates, la ecuación 3-57 se usa como el estadístico de prueba. Por la ecuación 3-58 se encuentra

$$\begin{aligned} S^2 &= \frac{1}{N-1}\left[\sum_{i=1}^{a}\sum_{j=1}^{n_i} R_{ij}^2 - \frac{N(N+1)^2}{4}\right] \\ &= \frac{1}{24}\left[5497.79 - \frac{25(26)^2}{4}\right] \\ &= 53.03 \end{aligned}$$

Tabla 3-14 Datos y rangos para el experimento de la resistencia a la tensión del ejemplo 3-1

Peso porcentual del algodón									
15		20		25		30		35	
y_{1j}	R_{1j}	y_{2j}	R_{2j}	y_{3j}	R_{3j}	y_{4j}	R_{4j}	y_{5j}	R_{5j}
7	2.0	12	9.5	14	11.0	19	20.5	7	2.0
7	2.0	17	14.0	18	16.5	25	25.0	10	5.0
15	12.5	12	9.5	18	16.5	22	23.0	11	7.0
11	7.0	18	16.5	19	20.5	19	20.5	15	12.5
9	4.0	18	16.5	19	20.5	23	24.0	11	7.0
$R_{i.}$	27.5		66.0		85.0		113.0		33.5

y el estadístico de prueba es

$$H = \frac{1}{S^2}\left[\sum_{i=1}^{a}\frac{R_{i.}^2}{n_i} - \frac{N(N+1)^2}{4}\right]$$
$$= \frac{1}{53.03}\left[5245.0 - \frac{25(26)^2}{4}\right]$$
$$= 19.25$$

Puesto que $H > \chi^2_{0.01,4}$, 4 = 13.28, se rechazaría la hipótesis nula y se concluiría que los tratamientos difieren. (El valor P para $H = 19.25$ es $P = 0.0002$.) Se trata de la misma conclusión obtenida por el análisis de varianza usual con la prueba F.

3-10.2 Comentarios generales sobre la transformación de rangos

Al procedimiento utilizado en la sección anterior de reemplazar las observaciones con sus rangos se le llama la **transformación de rangos.** Es una técnica muy poderosa y útil. Si se aplicara la prueba F común a los rangos en lugar de a los datos originales, se obtendría

$$F_0 = \frac{H/(a-1)}{(N-1-H)/(N-a)} \tag{3-60}$$

como el estadístico de prueba (ver Conover [20], p. 337). Observe que cuando el estadístico H de Kruskal-Wallis se incrementa o decrementa, F_0 también se incrementa o decrementa, por lo que la prueba de Kruskal-Wallis es equivalente a aplicar el análisis de varianza común a los rangos.

La transformación de rangos tiene una amplia aplicabilidad en los problemas de diseño experimental para los que no existe ninguna alternativa no paramétrica para el análisis de varianza. Esto incluye muchos de los diseños de capítulos subsecuentes de este libro. Si los datos están en rangos y se aplica la prueba F común, el resultado es un procedimiento aproximado que tiene buenas propiedades estadísticas (ver Conover e Iman [30a, b]). Cuando existe preocupación acerca del supuesto de normalidad o por el efecto de puntos atípicos o valores "absurdos", se recomienda que el análisis de varianza común se realice tanto en los datos originales como en los rangos. Cuando ambos procedimientos producen resultados similares, probablemente los supuestos del análisis de varianza se satisfacen razonablemente, y el análisis estándar es satisfactorio. Cuando los dos procedimientos difieren, deberá darse preferencia a la transformación de rangos, ya que es menos posible que sea distorsionada por una condición de no normalidad o la presencia de observaciones inusuales. En tales casos, tal vez el experimentador quiera investigar el uso de transfor-

maciones para la falta de normalidad y examinar los datos y el procedimiento experimental a fin de determinar si hay puntos atípicos y por qué han ocurrido.

3-11 PROBLEMAS

3-1. Se estudia la resistencia a la tensión del cemento portland. Pueden usarse económicamente cuatro diferentes técnicas de mezclado. Se han colectado los siguientes datos:

Técnica de mezclado	Resistencia a la tensión (lb/pulg²)			
1	3129	3000	2865	2890
2	3200	3300	2975	3150
3	2800	2900	2985	3050
4	2600	2700	2600	2765

a) Probar la hipótesis de que las técnicas de mezclado afectan la resistencia del cemento. Utilizar $\alpha = 0.05$.

b) Construir una representación gráfica como se describió en la sección 3-5.3 para comparar las resistencias a la tensión promedio de las cuatro técnicas de mezclado. ¿A qué conclusiones se llega?

c) Usar el método LSD de Fisher con $\alpha = 0.05$ para hacer comparaciones entre pares de medias.

d) Construir una gráfica de probabilidad normal de los residuales. ¿Qué conclusiones se sacarían acerca de la validez del supuesto de normalidad?

e) Graficar los residuales contra la resistencia a la tensión predicha. Comentar la gráfica.

f) Hacer un diagrama de dispersión de los resultados como ayuda para la interpretación de los resultados de este experimento.

3-2. *a*) Resolver de nuevo el inciso *b* del problema 3-1 utilizando la prueba del rango múltiple de Duncan con $\alpha = 0.05$. ¿Hay alguna diferencia en las conclusiones?

b) Resolver de nuevo el inciso *b* del problema 3-1 utilizando la prueba de Tukey con $\alpha = 0.05$. ¿Se llega a las mismas conclusiones con la prueba de Tukey que las obtenidas con el procedimiento gráfico y/o con la prueba del rango múltiple de Duncan?

c) Explicar la diferencia entre los procedimientos de Duncan y de Tukey.

3-3. Considere nuevamente el problema 3-1. Encontrar un intervalo de confianza de 95% para la resistencia a la tensión media del cemento portland que produce cada una de las cuatro técnicas de mezclado. Encontrar también un intervalo de confianza de 95% para la diferencia en las medias de las técnicas 1 y 3. ¿Sirve esto de ayuda para interpretar los resultados del experimento?

3-4. Se llevó a cabo un experimento a fin de determinar si cuatro temperaturas de cocción específicas afectan la densidad de cierto tipo de ladrillo. El experimento produjo los siguientes datos:

Temperatura	Densidad				
100	21.8	21.9	21.7	21.6	21.7
125	21.7	21.4	21.5	21.4	
150	21.9	21.8	21.8	21.6	21.5
175	21.9	21.7	21.8	21.4	

a) ¿La temperatura de cocción afecta la densidad de los ladrillos? Utilizar $\alpha = 0.05$.

b) ¿Es apropiado comparar las medias utilizando la prueba del rango múltiple de Duncan (por ejemplo) en este experimento?

c) Analizar los residuales de este experimento. ¿Se satisfacen los supuestos del análisis de varianza?

d) Construir una representación gráfica de los tratamientos como se describió en la sección 3-5.3. ¿Esta gráfica resume adecuadamente los resultados del análisis de varianza del inciso *a*?

3-5. Resolver de nuevo el inciso *d* del problema 3-4 utilizando el método LSD de Fisher. ¿A qué conclusiones se llega? Explicar en detalle cómo se modificó la técnica para tomar en cuenta los tamaños de las muestras desiguales.

3-6. Un fabricante de televisores está interesado en el efecto de cuatro tipos diferentes de recubrimientos para cinescopios de color sobre la conductividad de un cinescopio. Se obtienen los siguientes datos de la conductividad:

Tipo de recubrimiento	Conductividad			
1	143	141	150	146
2	152	149	137	143
3	134	136	132	127
4	129	127	132	129

a) ¿Hay alguna diferencia en la conductividad debida al tipo de recubrimiento? Utilizar $\alpha = 0.05$.
b) Estimar la media global y los efectos de los tratamientos.
c) Calcular la estimación de un intervalo de confianza de 95% para la media del tipo de recubrimiento 4. Calcular la estimación de un intervalo de confianza de 99% para la diferencia media entre los tipos de recubrimiento 1 y 4.
d) Probar todos los pares de medias utilizando el método LSD de Fisher con $\alpha = 0.05$.
e) Usar el método gráfico comentado en la sección 3-5.3 para comparar las medias. ¿Cuál es el tipo de recubrimiento que produce la conductividad más alta?
f) Suponiendo que el recubrimiento tipo 4 es el que se está usando actualmente, ¿qué se recomendaría al fabricante? Quiere minimizarse la conductividad.

3-7. Considere nuevamente el experimento del problema 3-6. Analizar los residuales y sacar conclusiones acerca de la adecuación del modelo.

3-8. En un artículo de *ACI Materials Journal* (vol. 84, pp. 213-216) se describen varios experimentos para investigar el varillado del concreto para eliminar el aire atrapado. Se usó un cilindro de 3 × 6 pulgadas; y el número de veces que esta barra se utilizó es la variable del diseño. La resistencia a la compresión resultante de la muestra de concreto es la respuesta. Los datos se muestran en la tabla siguiente:

Nivel de varillado	Resistencia a la compresión		
10	1530	1530	1440
15	1610	1650	1500
20	1560	1730	1530
25	1500	1490	1510

a) ¿Hay alguna diferencia en la resistencia a la compresión debida al nivel de varillado? Utilizar $\alpha = 0.05$.
b) Encontrar el valor *P* para el estadístico *F* del inciso *a*.
c) Analizar los residuales de este experimento. ¿Qué conclusiones pueden sacarse acerca de los supuestos fundamentales del modelo?
d) Construir una representación gráfica para comparar las medias de los tratamientos, como se describió en la sección 3-5.3.

3-9. En un artículo de *Environment International* (vol. 18, no. 4) se describe un experimento en el que se investigó la cantidad de radón liberado en las duchas. Se usó agua enriquecida con radón en el experimento, y se probaron seis diámetros diferentes de los orificios de las regaderas. Los datos del experimento se muestran en la siguiente tabla:

Diámetro de los orificios	Radón liberado (%)			
0.37	80	83	83	85
0.51	75	75	79	79
0.71	74	73	76	77
1.02	67	72	74	74
1.40	62	62	67	69
1.99	60	61	64	66

a) ¿El tamaño de los orificios afecta el porcentaje promedio del radón liberado? Utilizar $\alpha = 0.05$.
b) Encontrar el valor *P* para el estadístico *F* del inciso *a*.
c) Analizar los residuales de este experimento.
d) Encontrar un intervalo de confianza de 95% para el porcentaje promedio de radón liberado cuando el diámetro de los orificios es 1.40.
e) Construir una representación gráfica para comparar las medias de los tratamientos, como se describió en la sección 3-5.3. ¿Qué conclusiones pueden sacarse?

3-10. Se determinó el tiempo de respuesta en milisegundos para tres diferentes tipos de circuitos que podrían usarse en un mecanismo de desconexión automática. Los resultados se muestran en la siguiente tabla:

Tipo de circuito	Tiempo de respuesta				
1	9	12	10	8	15
2	20	21	23	17	30
3	6	5	8	16	7

a) Probar la hipótesis de que los tres tipos de circuitos tienen el mismo tiempo de respuesta. Utilizar $\alpha = 0.01$.
b) Usar la prueba de Tukey para comparar pares de medias de los tratamientos. Utilizar $\alpha = 0.01$.
c) Usar el procedimiento gráfico de la sección 3-5.3 para comparar las medias de los tratamientos. ¿Qué conclusiones pueden sacarse? ¿Cómo se comparan con las conclusiones del inciso *b*?
d) Construir un conjunto de contrastes ortogonales, suponiendo que al principio del experimento se sospechaba que el tiempo de respuesta del circuito tipo 2 era diferente del de los otros dos.
e) Si el lector fuera el ingeniero de diseño y quisiera minimizar el tiempo de respuesta, ¿qué tipo de circuito seleccionaría?
f) Analizar los residuales de este experimento. ¿Se satisfacen los supuestos del análisis de varianza básico?

3-11. Se estudia la vida efectiva de los fluidos aislantes en una carga acelerada de 35 kV. Se han obtenido datos de una prueba para cuatro tipos de fluidos. Los resultados fueron los siguientes:

Tipo de fluido	Vida (en horas) con 35 kV de carga					
1	17.6	18.9	16.3	17.4	20.1	21.6
2	16.9	15.3	18.6	17.1	19.5	20.3
3	21.4	23.6	19.4	18.5	20.5	22.3
4	19.3	21.1	16.9	17.5	18.3	19.8

a) ¿Hay algún indicio de que los fluidos difieran? Utilizar $\alpha = 0.05$.
b) ¿Cuál fluido seleccionaría el lector, dado que el objetivo es conseguir la vida efectiva más larga?
c) Analizar los residuales de este experimento. ¿Se satisfacen los supuestos del análisis de varianza básico?

3-12. Se estudian cuatro diferentes tipos de diseños de un circuito digital de computadora para comparar la cantidad de ruido presente. Se obtienen los siguientes datos:

Diseño del circuito	Ruido observado				
1	19	20	19	30	8
2	80	61	73	56	80
3	47	26	25	35	50
4	95	46	83	78	97

a) ¿La cantidad de ruido presente es la misma para los cuatro diseños? Utilizar $\alpha = 0.05$.
b) Analizar los residuales de este experimento. ¿Se satisfacen los supuestos del análisis de varianza?
c) ¿Qué diseño del circuito se seleccionaría para usarlo? El ruido bajo es mejor.

3-13. Se pide a cuatro químicos que determinen el porcentaje de alcohol metílico en cierto compuesto químico. Cada químico hace tres determinaciones, y los resultados son los siguientes:

Químico	Porcentaje de alcohol metílico		
1	84.99	84.04	84.38
2	85.15	85.13	84.88
3	84.72	84.48	85.16
4	84.20	84.10	84.55

a) ¿Los químicos difieren significativamente? Utilizar $\alpha = 0.05$.
b) Analizar los residuales de este experimento.
c) Si el químico 2 es un empleado nuevo, construir un conjunto razonable de contrastes ortogonales que podría haberse usado al principio del experimento.

3-14. Se someten a estudio tres marcas de baterías. Se sospecha que las vidas (en semanas) de las tres marcas son diferentes. Se prueban cinco baterías de cada marca con los resultados siguientes:

Semanas de vida		
Marca 1	Marca 2	Marca 3
100	76	108
96	80	100
92	75	96
96	84	98
92	82	100

a) ¿Las vidas de estas tres marcas son diferentes?
b) Analizar los residuales de este experimento.
c) Construir la estimación de un intervalo de confianza de 95% para la vida media de la batería marca 2. Construir la estimación del intervalo de confianza de 99% para la diferencia media entre las vidas de las baterías marcas 2 y 3.

d) ¿Qué marca seleccionaría el lector para usarla? Si el fabricante reemplazara sin cargo cualquier batería que dure menos de 85 semanas, ¿qué porcentaje esperaría reemplazar la compañía?

3-15. Se están investigando cuatro catalizadores que pueden afectar la concentración de un componente en una mezcla líquida de tres componentes. Se obtienen las siguientes concentraciones:

Catalizador			
1	2	3	4
58.2	56.3	50.1	52.9
57.2	54.5	54.2	49.9
58.4	57.0	55.4	50.0
55.8	55.3		51.7
54.9			

a) ¿Los cuatro catalizadores tienen el mismo efecto sobre la concentración?
b) Analizar los residuales de este experimento.
c) Construir la estimación de un intervalo de confianza de 99% para la respuesta media del catalizador 1.

3-16. Se llevó a cabo un experimento para investigar la eficacia de cinco materiales aislantes. Se probaron cuatro muestras de cada material con un nivel elevado de voltaje para acelerar el tiempo de falla. Los tiempos de falla (en minutos) se muestran abajo:

Material	Tiempo de falla (minutos)			
1	110	157	194	178
2	1	2	4	18
3	880	1256	5276	4355
4	495	7040	5307	10,050
5	7	5	29	2

a) ¿Los cinco materiales tienen el mismo efecto sobre el tiempo de falla?
b) Graficar los residuales contra la respuesta predicha. Construir una gráfica de probabilidad normal de los residuales. ¿Qué información transmiten estas gráficas?
c) Con base en la respuesta del inciso *b*, realizar otro análisis de los datos del tiempo de falla y sacar las conclusiones apropiadas.

3-17. Un fabricante de semiconductores ha desarrollado tres métodos diferentes para reducir el conteo de partículas en las obleas. Los tres métodos se prueban en cinco obleas y se obtiene el conteo de partículas después del tratamiento. Los datos se muestran abajo:

Método	Conteo				
1	31	10	21	4	1
2	62	40	24	30	35
3	53	27	120	97	68

a) ¿Todos los métodos tienen el mismo efecto sobre el conteo promedio de partículas?
b) Graficar los residuales contra la respuesta predicha. Construir una gráfica de probabilidad normal de los residuales. ¿Hay motivo de preocupación potencial acerca de la validez de los supuestos?
c) Con base en la respuesta del inciso *b*, realizar otro análisis de los datos del conteo de partículas y sacar las conclusiones apropiadas.

3-18. Considere la prueba de la igualdad de las medias de dos poblaciones normales, donde las varianzas son desconocidas pero se suponen iguales. El procedimiento de prueba apropiado es la prueba t agrupada o combinada. Demostrar que la prueba t combinada es equivalente al análisis de varianza de un solo factor.

3-19. Demostrar que la varianza de la combinación lineal $\Sigma_{i=1}^{a} c_i y_{i.}$ es $\sigma^2 \Sigma_{i=1}^{a} n_i c_i^2$.

3-20. En un experimento con efectos fijos, suponga que hay n observaciones para cada uno de cuatro tratamientos. Sean Q_1^2, Q_2^2, Q_3^2 los componentes con un solo grado de libertad de los contrastes ortogonales. Demostrar que $SS_{\text{Tratamientos}} = Q_1^2 + Q_2^2 + Q_3^2$.

3-21. Utilizar la prueba de Bartlett para determinar si el supuesto de la igualdad de las varianzas se satisface en el problema 3-14. Utilizar $\alpha = 0.05$. ¿Se llegó a la misma conclusión respecto de la igualdad de las varianzas con el examen de las gráficas de los residuales?

3-22. Utilizar la prueba de Levene modificada para determinar si el supuesto de las varianzas iguales se satisface en el problema 3-14. Utilizar $\alpha = 0.05$. ¿Se llegó a la misma conclusión respecto de la igualdad de las varianzas con el examen de las gráficas de los residuales?

3-23. Referirse al problema 3-10. Si quiere detectarse una diferencia máxima en los tiempos de respuesta promedio de 10 milisegundos con una probabilidad de al menos 0.90, ¿qué tamaño de la muestra deberá usarse? ¿Cómo se obtendría una estimación preliminar de σ^2?

3-24. Referirse al problema 3-14.

a) Si quiere detectarse una diferencia máxima en la vida de las baterías de 10 horas con una probabilidad de al menos 0.90, ¿qué tamaño de la muestra deberá usarse? Comentar cómo se obtendría una estimación preliminar de σ^2 para responder esta pregunta.

b) Si la diferencia entre las marcas es lo suficientemente grande para que la desviación estándar de una observación se incremente en 25%, ¿qué tamaño de la muestra deberá usarse si quiere detectarse esto con una probabilidad de al menos 0.90?

3-25. Considere el experimento del problema 3-14. Si quiere construirse un intervalo de confianza de 95% para la diferencia en las vidas medias de dos baterías que tenga una precisión de ±2 semanas, ¿cuántas baterías de cada marca deben probarse?

3-26. Suponga que cuatro poblaciones normales tienen medias $\mu_1 = 50, \mu_2 = 60, \mu_3 = 50$ y $\mu_4 = 60$. ¿Cuántas observaciones deberán hacerse en cada población para que la probabilidad de rechazar la hipótesis nula de la igualdad de las medias poblacionales sea al menos 0.90? Suponer que $\alpha = 0.05$ y que una estimación razonable de la varianza de error es $\sigma^2 = 25$.

3-27. Referirse al problema 3-26.

a) ¿En qué forma cambiaría la respuesta si una estimación razonable de la varianza del error experimental fuera $\sigma^2 = 36$?

b) ¿En qué forma cambiaría la respuesta si una estimación razonable de la varianza del error experimental fuera $\sigma^2 = 49$?

c) ¿Puede sacarse alguna conclusión acerca de la sensibilidad de la respuesta dada en esta situación particular acerca de cómo afecta la estimación de σ la decisión referente al tamaño de la muestra?

d) ¿Puede hacerse alguna recomendación acerca de cómo debería usarse este enfoque general para elegir n en la práctica?

3-28. Referirse al experimento de la fundición de aluminio descrito en la sección 3-8. Verificar que los métodos para controlar la proporción de alúmina no afectan el voltaje promedio de la celda. Construir una gráfica de probabilidad normal de los residuales. Graficar los residuales contra los valores predichos. ¿Existe algún indicio de que se violan algunos de los supuestos fundamentales?

3-29. Referirse al experimento de la fundición de aluminio de la sección 3-8. Verificar el análisis de varianza del ruido del crisol que se resume en la tabla 3-13. Examinar las gráficas de los residuales usuales y comentar la validez del experimento.

3-30. Se investigaron cuatro diferentes velocidades de alimentación en un experimento con una máquina CNC que produce una pieza que se usa en la unidad de potencia auxiliar de un avión. El ingeniero de manufactura a cargo del experimento sabe que una dimensión crítica de la pieza de interés puede ser afectada por la velocidad de alimentación. Sin embargo, la experiencia previa indica que es probable que sólo estén presentes

efectos de dispersión. Es decir, al cambiarse la velocidad de alimentación no se afecta la dimensión *promedio*, pero podría afectarse la variabilidad dimensional. El ingeniero realiza cinco corridas de producción con cada velocidad de alimentación y obtiene la desviación estándar de la dimensión crítica (en 10^{-3} mm). Los datos se muestran abajo. Suponer que todas las corridas se hicieron en orden aleatorio.

Velocidad de alimentación (pulgadas/minuto)	Corrida de producción				
	1	2	3	4	5
10	0.09	0.10	0.13	0.08	0.07
12	0.06	0.09	0.12	0.07	0.12
14	0.11	0.08	0.08	0.05	0.06
16	0.19	0.13	0.15	0.20	0.11

a) ¿La velocidad de alimentación tiene algún efecto sobre la desviación estándar de esta dimensión crítica?

b) Usar los residuales de este experimento para investigar la adecuación del modelo. ¿Hay algún problema con la validez experimental?

3-31. Considere los datos del problema 3-10.

a) Escribir las ecuaciones normales de mínimos cuadrados para este problema y resolverlas para $\hat{\mu}$ y $\hat{\tau}_i$, utilizando la restricción usual ($\Sigma_{i=1}^{3}\hat{\tau}_i = 0$). Estimar $\tau_1 - \tau_2$.

b) Resolver las ecuaciones del inciso *a* utilizando la restricción $\hat{\tau}_3 = 0$. ¿Los estimadores $\hat{\tau}_i$ y $\hat{\mu}$ son los mismos que se encontraron en el inciso *a*? ¿Por qué? Estimar ahora $\tau_1 - \tau_2$ y comparar la respuesta con la del inciso *a*. ¿Qué afirmación puede hacerse respecto de estimar los contrastes en las τ_i?

c) Estimar $\mu + \tau_1$, $2\tau_1 - \tau_2 - \tau_3$, y $\mu + \tau_1 + \tau_2$ utilizando las dos soluciones de las ecuaciones normales. Comparar los resultados obtenidos en cada caso.

3-32. Aplicar la prueba general de significación de la regresión en el experimento del ejemplo 3-1. Demostrar que el procedimiento produce los mismos resultados que el análisis de varianza usual.

3-33. Usar la prueba de Kruskal-Wallis en el experimento del problema 3-11. Comparar las conclusiones obtenidas con las del análisis de varianza usual.

3-34. Usar la prueba de Kruskal-Wallis en el experimento del problema 3-12. ¿Los resultados son comparables con los encontrados por el análisis de varianza usual?

3-35. Considere el experimento del ejemplo 3-1. Suponga que la observación mayor de la resistencia a la tensión se registró incorrectamente como 50. ¿Qué efecto tiene esto sobre el análisis de varianza usual? ¿Qué efecto tiene sobre la prueba de Kruskal-Wallis?

4 Bloques aleatorizados, cuadrados latinos y diseños relacionados

4-1 DISEÑO DE BLOQUES COMPLETOS ALEATORIZADOS

En cualquier experimento, la variabilidad que surge de un factor perturbador puede afectar los resultados. En general, un **factor perturbador** puede definirse como un factor del diseño que probablemente tenga un efecto sobre la respuesta, pero en el que no existe un interés específico. En ocasiones un factor perturbador es **desconocido y no controlable**; es decir, se desconoce la existencia de ese factor e incluso puede tener niveles variables mientras se está realizando el experimento. La **aleatorización** es la técnica de diseño que se utiliza para protegerse contra estos factores perturbadores "que están al acecho". En otros casos, el factor perturbador es **conocido pero no controlable.** Si por lo menos puede observarse el valor que asume el factor perturbador en cada corrida del experimento, es posible hacer la compensación correspondiente en el análisis estadístico mediante el uso del **análisis de covarianza,** una técnica que se revisará en el capítulo 14. Cuando la fuente de variabilidad perturbadora es **conocida y controlable,** puede usarse una técnica de diseño llamada **formación de bloques** para eliminar de manera sistemática su efecto sobre las comparaciones estadísticas entre los tratamientos. La formación de bloques es una técnica de diseño en extremo importante que se utiliza ampliamente en la experimentación industrial, y es la materia de este capítulo.

Para ilustrar la idea general, suponga que quiere determinarse si cuatro puntas diferentes producen o no lecturas diferentes en una máquina para probar la dureza. Un experimento como éste podría ser parte de un estudio de la aptitud en la calibración de los instrumentos. La máquina funciona presionando la punta en un ejemplar de prueba de metal, y por la profundidad de la depresión resultante puede determinarse la dureza del ejemplar. El experimentador ha decidido obtener cuatro observaciones para cada punta. Hay un solo factor —el tipo de punta—, y un diseño completamente aleatorizado de un solo factor consistiría en asignar al azar cada una de las $4 \times 4 = 16$ corridas a una **unidad experimental,** es decir, a un ejemplar de prueba de metal, y observar qué resulta de la lectura de la dureza. Por lo tanto, se necesitarían 16 ejemplares de prueba de metal en este experimento, uno por cada corrida del diseño.

Existe un problema potencialmente serio con un experimento por completo aleatorizado en esta situación de diseño. Si los ejemplares de prueba de metal difieren ligeramente en sus durezas, como podría

Tabla 4-1 Diseño de bloques completos aleatorizados para el experimento de la prueba de la dureza

Tipo de punta	Ejemplar de prueba			
	1	2	3	4
1	9.3	9.4	9.6	10.0
2	9.4	9.3	9.8	9.9
3	9.2	9.4	9.5	9.7
4	9.7	9.6	10.0	10.2

ocurrir si se tomaran de lingotes que se produjeron con temperaturas diferentes, las unidades experimentales (los ejemplares de prueba) contribuirán a la variabilidad observada en los datos de la dureza. Como resultado, el error experimental reflejará *tanto* el error aleatorio *como* la variabilidad entre los ejemplares de prueba.

El objetivo sería hacer el error experimental tan pequeño como fuera posible; es decir, querría eliminarse del error experimental la variabilidad entre los ejemplares de prueba. Un diseño para lograr esto requiere que el experimentador pruebe cada punta una vez en cada uno de los cuatro ejemplares de prueba. A este diseño, que se muestra en la tabla 4-1, se le llama **diseño de bloques completos aleatorizados (RCBD, *randomized complete block design*)**. La respuesta observada es la dureza en la escala C de Rockwell menos 40. La palabra "completos" indica que cada bloque (ejemplar de prueba) contiene todos los tratamientos (puntas). Al utilizar este diseño, los bloques o ejemplares de prueba forman una unidad experimental más homogénea en la cual comparar las puntas. De hecho, esta estrategia de diseño mejora la precisión de las comparaciones entre las puntas al eliminar la variabilidad entre los ejemplares de prueba. Dentro de un bloque, el orden en que se prueban las cuatro puntas se determina aleatoriamente. Observe la similitud de este problema de diseño con el de la sección 2-5, donde se analizó la prueba *t* pareada. El diseño de bloques completos aleatorizados es una generalización de ese concepto.

El RCBD es uno de los diseños experimentales más utilizados. Son numerosas las situaciones en las que el RCBD es apropiado. Las unidades de equipo o maquinaria de prueba son con frecuencia diferentes en sus características de operación y serían un factor de formación de bloques típico. Lotes de materia prima, personas y el tiempo también son fuentes de variabilidad perturbadora comunes en un experimento que pueden controlarse de manera sistemática mediante la formación de bloques.

La formación de bloques también puede ser útil en situaciones que no incluyen necesariamente factores perturbadores. Por ejemplo, suponga que un ingeniero químico está interesado en el efecto de la velocidad de alimentación del catalizador sobre la viscosidad de un polímero. Sabe que hay varios factores, como la fuente de la materia prima, la temperatura, el operador y la pureza de la materia prima, que son muy difíciles de controlar en proceso en gran escala. Por lo tanto, decide probar en bloques la velocidad de alimentación del catalizador, donde cada bloque consiste en alguna combinación de estos factores no controlables. De hecho, está utilizando los bloques para probar la **robustez** de su variable de proceso (la velocidad de alimentación) para las condiciones que no puede controlar con facilidad. Para un análisis más amplio de este punto, ver Coleman y Montgomery [27].

4-1.1 Análisis estadístico del diseño de bloques completos aleatorizados

Suponga que se tienen, en general, a tratamientos que van a compararse y b bloques. El diseño de bloques completos aleatorizados se muestra en la figura 4-1. Hay una observación por tratamiento en cada bloque, y el orden en que se corren los tratamientos dentro de cada bloque se determina al azar. Debido a

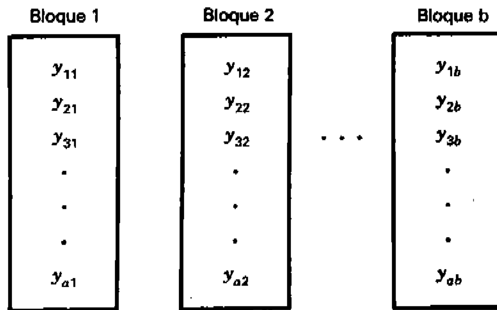


Figura 4-1 El diseño de bloques completos aleatorizados.

que la única aleatorización de los tratamientos se hace dentro de los bloques, con frecuencia se dice que los bloques representan una **restricción sobre la aleatorización.**

El **modelo estadístico** del RCBD puede escribirse de varias maneras. El tradicional es el **modelo de los efectos**:

$$y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij} \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, b \end{cases} \tag{4-1}$$

donde μ es la media global, τ_i es el efecto del tratamiento i-ésimo, β_j es el efecto del bloque j-ésimo, y ε_{ij} es el término del error NID$(0, \sigma^2)$ usual. Se considerará inicialmente que los tratamientos y los bloques son factores fijos. Como en el modelo del diseño experimental con un solo factor del capítulo 3, el modelo de los efectos para el RCBD es un modelo sobreespecificado. En consecuencia, los efectos de los tratamientos y los bloques se consideran por lo general como desviaciones de la media global, por lo que

$$\sum_{i=1}^{a} \tau_i = 0 \quad \text{y} \quad \sum_{j=1}^{b} \beta_j = 0$$

También es posible usar un **modelo de las medias** para el RCBD, por ejemplo

$$y_{ij} = \mu_{ij} + \varepsilon_{ij} \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, b \end{cases}$$

donde $\mu_{ij} = \mu + \tau_i + \beta_j$. Sin embargo, en este capítulo se usará el modelo de los efectos de la ecuación 4-1.

En un experimento en el que se use el RCBD, el interés se encuentra en probar la igualdad de las medias de los tratamientos. Por lo tanto, las hipótesis de interés son

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_a$$
$$H_1 : \text{al menos una } \mu_i \neq \mu_j$$

Puesto que la media del tratamiento i-ésimo es $\mu_i = (1/b)\Sigma_{j=1}^{b}(\mu + \tau_i + \beta_j) = \mu + \tau_i$, una manera equivalente de escribir las hipótesis anteriores es en términos de los efectos de los tratamientos, por ejemplo

$$H_0 : \tau_1 = \tau_2 = \cdots = \tau_a = 0$$
$$H_1 : \tau_i \neq 0 \text{ para al menos una } i$$

Sea $y_{i.}$ el total de observaciones hechas bajo el tratamiento i, $y_{.j}$ el total de observaciones del bloque j, $y_{..}$ el gran total de las observaciones y $N = ab$ el número total de observaciones. Expresado matemáticamente,

$$y_{i.} = \sum_{j=1}^{b} y_{ij} \qquad i = 1, 2, \ldots, a \tag{4-2}$$

$$y_{.j} = \sum_{i=1}^{a} y_{ij} \qquad j = 1, 2, \ldots, b \tag{4-3}$$

y

$$y_{..} = \sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij} = \sum_{i=1}^{a} y_{i.} = \sum_{j=1}^{b} y_{.j} \tag{4-4}$$

De manera similar, $\bar{y}_{i.}$ es el promedio de las observaciones hechas bajo el tratamiento i, $\bar{y}_{.j}$ es el promedio de las observaciones del bloque j, $\bar{y}_{..}$ es el gran promedio de todas las observaciones. Es decir,

$$\bar{y}_{i.} = y_{i.} / b \qquad \bar{y}_{.j} = y_{.j} / a \qquad \bar{y}_{_} / N \tag{4-5}$$

La suma de cuadrados total corregida puede expresarse como

$$\sum_{i=1}^{a}\sum_{j=1}^{b} (y_{ij} - \bar{y}_{..})^2 = \sum_{i=1}^{a}\sum_{j=1}^{b} [(\bar{y}_{i.} - \bar{y}_{..}) + (\bar{y}_{.j} - \bar{y}_{..}) + (y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})]^2 \tag{4-6}$$

Al desarrollar el miembro del lado derecho de la ecuación 4-6 se obtiene

$$\begin{aligned}\sum_{i=1}^{a}\sum_{j=1}^{b} (y_{ij} - \bar{y}_{..})^2 = {} & b\sum_{i=1}^{a} (\bar{y}_{i.} - \bar{y}_{..})^2 + a\sum_{j=1}^{b} (\bar{y}_{.j} - \bar{y}_{..})^2 \\ & + \sum_{i=1}^{a}\sum_{j=1}^{b} (y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})^2 + 2\sum_{i=1}^{a}\sum_{j=1}^{b} (\bar{y}_{i.} - \bar{y}_{..})(\bar{y}_{.j} - \bar{y}_{..}) \\ & + 2\sum_{i=1}^{a}\sum_{j=1}^{b} (\bar{y}_{.j} - \bar{y}_{..})(y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..}) \\ & + 2\sum_{i=1}^{a}\sum_{j=1}^{b} (\bar{y}_{i.} - \bar{y}_{..})(y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})\end{aligned}$$

Mediante procedimientos algebraicos simples, pero laboriosos, se prueba que los tres productos cruzados son cero. Por lo tanto,

$$\begin{aligned}\sum_{i=1}^{a}\sum_{j=1}^{b} (y_{ij} - \bar{y}_{..})^2 = {} & b\sum_{i=1}^{a} (\bar{y}_{i.} - \bar{y}_{..})^2 + a\sum_{j=1}^{b} (\bar{y}_{.j} - \bar{y}_{..})^2 \\ & + \sum_{i=1}^{a}\sum_{j=1}^{b} (y_{ij} - \bar{y}_{.j} - \bar{y}_{i.} + \bar{y}_{..})^2\end{aligned} \tag{4-7}$$

representa una partición de la suma de cuadrados total. Al expresar simbólicamente las sumas de cuadrados de la ecuación 4-7, se tiene

$$SS_T = SS_{\text{Tratamientos}} + SS_{\text{Bloques}} + SS_E \tag{4-8}$$

Puesto que hay N observaciones, SS_T tiene $N - 1$ grados de libertad. Hay a tratamientos y b bloques, de donde $SS_{\text{Tratamientos}}$ y SS_{Bloques} tienen $a - 1$ y $b - 1$ grados de libertad, respectivamente. La suma de cuadra-

dos del error es sólo la suma de cuadrados entre las celdas menos la suma de cuadrados de los tratamientos y los bloques. Hay ab celdas con $ab-1$ grados de libertad entre ellas, de donde SS_E tiene $ab-1-(a-1)-(b-1)=(a-1)(b-1)$ grados de libertad. Además, la suma de los grados de libertad del lado derecho de la ecuación 4-8 es igual al total del lado izquierdo; por lo tanto, al establecer los supuestos de normalidad usuales para los errores, puede usarse el teorema 3-1 para demostrar que $SS_{\text{Tratamientos}}/\sigma^2$, $SS_{\text{Bloques}}/\sigma^2$ y SS_E/σ^2 son variables aleatorias ji-cuadrada con distribuciones independientes. Cada suma de cuadrados dividida por sus grados de libertad es un cuadrado medio. Puede demostrarse que el valor esperado de los cuadrados medios, si los tratamientos y los bloques son fijos, es

$$E(MS_{\text{Tratamientos}})=\sigma^2+\frac{b\sum_{i=1}^{a}\tau_i^2}{a-1}$$

$$E(MS_{\text{Bloques}})=\sigma^2+\frac{a\sum_{j=1}^{b}\beta_j^2}{b-1}$$

$$E(MS_E)=\sigma^2$$

Por lo tanto, para probar la igualdad de las medias de los tratamientos, se usaría el estadístico de prueba

$$F_0=\frac{MS_{\text{Tratamientos}}}{MS_E}$$

que se distribuye como $F_{a-1,(a-1)(b-1)}$ si la hipótesis nula es verdadera. La región crítica es la cola superior de la distribución F, y H_0 se rechazaría si $F_0 > F_{\alpha,\, a-1,\, (a-1)(b-1)}$.

También podría haber interés en comparar las medias de los bloques porque, en caso de que la diferencia entre estas medias no sea considerable, quizá no sea necesaria la formación de bloques en experimentos futuros. Por los cuadrados medios esperados, aparentemente la hipótesis $H_0{:}\beta_j = 0$ puede probarse comparando el estadístico $F_0 = MS_{\text{Bloques}}/MS_E$ con $F_{\alpha,\, b-1,\, (a-1)(b-1)}$. Sin embargo, recuerde que la aleatorización sólo se ha aplicado a los tratamientos *dentro* de los bloques; es decir, los bloques representan una **restricción sobre la aleatorización.** ¿Qué efecto tiene esto sobre el estadístico $F_0 = MS_{\text{Bloques}}/MS_E$? Existen diferentes puntos de vista para abordar esta cuestión. Por ejemplo, Box, Hunter y Hunter [18] señalan que la prueba F del análisis de varianza común puede justificarse exclusivamente con base en la aleatorización,[1] sin el uso directo del supuesto de normalidad. Agregan que en la prueba para comparar las medias de los bloques no puede recurrirse a dicha justificación debido a la restricción sobre la aleatorización; pero si los errores son NID(0, σ^2), puede usarse el estadístico $F_0 = MS_{\text{Bloques}}/MS_E$ para comparar las medias de los bloques. Por otra parte, Anderson y McLean [2] argumentan que la restricción sobre la aleatorización impide que este estadístico sea una prueba significativa para comparar las medias de los bloques y que este cociente F es en realidad una prueba de la igualdad de las medias de los bloques más la restricción sobre la aleatorización (a la que llaman el error de la restricción; ver Anderson y McLean [2] para detalles adicionales).

Entonces, ¿qué se hace en la práctica? Debido a que con frecuencia el supuesto de normalidad es cuestionable, considerar $F_0 = MS_{\text{Bloques}}/MS_E$ como una prueba F exacta para la igualdad de las medias de los bloques no es una buena práctica general. Por esa razón, esta prueba F no se incluye en la tabla del análisis de varianza. Sin embargo, como un procedimiento aproximado para investigar el efecto de la variable formación de bloques, examinar el cociente MS_{Bloques}/MS_E es muy razonable. Si este cociente es muy

[1] De hecho, la distribución F de la teoría normal es una aproximación de la distribución de aleatorización generada al calcular F_0 a partir de cada asignación posible de las respuestas a los tratamientos.

Tabla 4-2 Análisis de varianza de un diseño de bloques completos aleatorizados

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0
Tratamientos	$SS_{\text{Tratamientos}}$	$a-1$	$\frac{SS_{\text{Tratamientos}}}{a-1}$	$\frac{MS_{\text{Tratamientos}}}{MS_E}$
Bloques	SS_{Bloques}	$b-1$	$\frac{SS_{\text{Bloques}}}{b-1}$	
Error	SS_E	$(a-1)(b-1)$	$\frac{SS_E}{(a-1)(b-1)}$	
Total	SS_T	$N-1$		

grande, implica que el factor formación de bloques tiene un efecto considerable y que la reducción del ruido obtenida por la formación de bloques probablemente fue útil para mejorar la precisión de la comparación de las medias de los tratamientos.

El procedimiento suele resumirse en un esquema de análisis de varianza, como el que se muestra en la tabla 4-2. En general, los cálculos se realizarían con un paquete de software de estadística. Sin embargo, es posible obtener fórmulas de cálculo manual de las sumas de cuadrados para los elementos de la ecuación 4-7 expresándolos en términos de los totales de los tratamientos y los bloques. Estas fórmulas de cálculo son

$$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij}^2 - \frac{y_{..}^2}{N} \tag{4-9}$$

$$SS_{\text{Tratamientos}} = \frac{1}{b}\sum_{i=1}^{a} y_{i.}^2 - \frac{y_{..}^2}{N} \tag{4-10}$$

$$SS_{\text{Bloques}} = \frac{1}{a}\sum_{j=1}^{b} y_{.j}^2 - \frac{y_{..}^2}{N} \tag{4-11}$$

y la suma de cuadrados del error se obtiene por sustracción como

$$SS_E = SS_T - SS_{\text{Tratamientos}} - SS_{\text{Bloques}} \tag{4-12}$$

EJEMPLO 4-1 ··

Considere el experimento de la prueba de la dureza de la sección 4-1. Hay cuatro puntas y cuatro ejemplares de prueba de metal. Cada punta se prueba una vez en cada ejemplar, resultando un diseño de bloques completos aleatorizados. Los datos obtenidos se repiten por conveniencia en la tabla 4-3. Recuerde que el *orden* en que se probaron las puntas en un ejemplar particular se determinó al azar. Para simplificar los

Tabla 4-3 Diseño de bloques completos aleatorizados para el experimento de la prueba de la dureza

	Ejemplar de prueba (bloque)			
Tipo de punta	1	2	3	4
1	9.3	9.4	9.6	10.0
2	9.4	9.3	9.8	9.9
3	9.2	9.4	9.5	9.7
4	9.7	9.6	10.0	10.2

Tabla 4-4 Datos codificados del experimento de la prueba de la dureza

	Ejemplar de prueba (bloque)				
Tipo de punta	1	2	3	4	$y_{i.}$
1	−2	−1	1	5	3
2	−1	−2	3	4	4
3	−3	−1	0	2	−2
4	2	1	5	7	15
$y_{.j}$	−4	−3	9	18	$20 = y_{..}$

cálculos, los datos originales se codifican restando 9.5 de cada observación y multiplicando el resultado por 10. Se obtienen así los datos de la tabla 4-4. Las sumas de cuadrados se obtienen de la siguiente manera:

$$SS_T = \sum_{i=1}^{4}\sum_{j=1}^{4} y_{ij}^2 - \frac{y_{..}^2}{N}$$

$$= 154.00 - \frac{(20)^2}{16} = 129.00$$

$$SS_{\text{Tratamientos}} = \frac{1}{b}\sum_{i=1}^{4} y_{i.}^2 - \frac{y_{..}^2}{N}$$

$$= \frac{1}{4}[(3)^2 + (4)^2 + (-2)^2 + (15)^2] - \frac{(20)^2}{16} = 38.50$$

$$SS_{\text{Bloques}} = \frac{1}{a}\sum_{j=1}^{4} y_{.j}^2 - \frac{y_{..}^2}{N}$$

$$= \frac{1}{4}[(-4)^2 + (-3)^2 + (9)^2 + (18)^2] - \frac{(20)^2}{16} = 82.50$$

$$SS_E = SS_T - SS_{\text{Tratamientos}} - SS_{\text{Bloques}}$$

$$= 129.00 - 38.50 - 82.50 = 8.00$$

En la tabla 4-5 se presenta el análisis de varianza. Utilizando $\alpha = 0.05$, el valor crítico de F es $F_{0.05,3,9} = 3.86$. Puesto que 14.44 > 3.86, se concluye que el tipo de punta afecta la lectura de la dureza media. El valor P para la prueba también es muy pequeño. Además, al parecer los ejemplares (bloques) difieren de manera significativa, ya que el cuadrado medio de los bloques es grande en relación con el error.

Es interesante observar los resultados que se habrían obtenido si no se hubiera tenido conocimiento de los diseños de bloques aleatorizados. Suponga que se usaran cuatro ejemplares, asignando al azar las puntas a cada uno de ellos, y que resultara (por casualidad) el mismo diseño que el de la tabla 4-3. El análisis incorrecto de estos datos como un diseño completamente aleatorizado de un solo factor se presenta en la tabla 4-6.

Tabla 4-5 Análisis de varianza del experimento de la prueba de la dureza

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0	Valor P
Tratamientos (tipo de punta)	38.50	3	12.83	14.44	0.0009
Bloques (ejemplares)	82.50	3	27.50		
Error	8.00	9	0.89		
Total	129.00	15			

Tabla 4-6 Análisis incorrecto del experimento de la prueba de la dureza como un diseño completamente aleatorizado

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0
Tipo de punta	38.50	3	12.83	1.70
Error	90.50	12	7.54	
Total	129.00	15		

Puesto que $F_{0.05,3,12} = 3.49$, no puede rechazarse la hipótesis de la igualdad de las mediciones de la dureza media de las cuatro puntas. Por lo tanto, el diseño de bloques aleatorizados reduce lo suficiente la cantidad de ruido en los datos para que las diferencias entre las cuatro puntas sean detectadas. Esto ilustra un punto muy importante. Si un experimentador no recurre a la formación de bloques cuando debería haberlo hecho, el efecto puede ser inflar el error experimental a tal grado que las diferencias importantes entre las medias de los tratamientos sean indetectables.

Muestra de salida de computadora

En la figura 4-2 se muestra la salida de computadora condensada obtenida con *Design-Expert* para los datos de la prueba de la dureza del ejemplo 4-1. Recuerde que en el análisis original de la tabla 4-5 se utilizaron datos *codificados*. (Las respuestas originales se codificaron restando 9.5 y multiplicando el resultado por 10.) El análisis de computadora utilizó las respuestas originales. Por consiguiente, las sumas de cuadrados de la figura 4-2 son iguales a las de la tabla 4-5 divididas entre 100 (observe que *Design-Expert* ha redondeado las sumas de cuadrados con dos cifras decimales).

Los residuales se enlistan en la parte inferior de la salida de computadora. Éstos se calculan como

$$e_{ij} = y_{ij} - \hat{y}_{ij}$$

y, como se demostrará más adelante, los valores ajustados son $\hat{y}_{ij} = \bar{y}_{i.} + \bar{y}_{.j} - \bar{y}_{..}$, de donde

$$e_{ij} = y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..} \tag{4-13}$$

En la sección siguiente se indicará cómo se usan los residuales en la **verificación de la adecuación del modelo.**

Comparaciones múltiples

Si los tratamientos en un RCBD son fijos, y el análisis indica una diferencia significativa en las medias de los tratamientos, al experimentador le interesarán por lo general comparaciones múltiples para descubrir *cuáles* son los tratamientos cuyas medias difieren. Para ello puede utilizarse cualquiera de los procedimientos de comparaciones múltiples del capítulo 3 (sección 3-5). Simplemente se sustituye en las fórmulas de la sección 3-5 el número de réplicas (n) en el diseño completamente aleatorizado de un solo factor con el número de bloques (b). Asimismo, es necesario recordar usar el número de grados de libertad del error para el bloque aleatorizado $[(a-1)(b-1)]$ en lugar de los grados de libertad del diseño completamente aleatorizado $[a(n-1)]$.

En la salida de *Design-Expert* de la figura 4-2 se ilustra el procedimiento LSD de Fisher. Observe que si se usa $\alpha = 0.05$, se concluiría que $\mu_2 = \mu_3$. Ahora bien, puesto que $\bar{y}_{3.} \leq \bar{y}_{1.} \leq \bar{y}_{2.}$ (es decir, las medias $\bar{y}_{2.}$ y $\bar{y}_{3.}$ abarcan algunas de las medias restantes), una conclusión inmediata sería que $\mu_1 = \mu_2 = \mu_3$. Además, μ_4 es diferente de las otras tres medias. Se concluye por lo tanto que la punta tipo 4 produce una dureza media que es significativamente más alta que las lecturas de la dureza media de los otros tres tipos de puntas.

Response: Hardness in Rockwell C
ANOVA for Selected Factorial Model
Analysis of variance table [Partial sum of squares]

Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Block	0.82	3	0.27			
Model	0.38	3	0.13	14.44	0.0009	significant
A	*0.38*	*3*	*0.13*	*14.44*	*0.0009*	
Residual	0.080	9	8.889E-003			
Cor Total	1.29	15				

Std. Dev.	0.094	R-Squared	0.8280
Mean	9.63	Adj R-Squared	0.7706
C.V.	0.98	Pred R-Squared	0.4563
PRESS	0.25	Adeq Precision	15.635

Treatment Means (Adjusted, If Necessary)

	Estimated Mean	Standard Error
1-A1	9.57	0.47
2-A2	9.60	0.47
3-A3	9.45	0.47
4-A4	9.88	0.47

Treatment	Mean Difference	DF	Standard Error	t for H_0 Coeff=0	Prob > |t|
1 vs 2	−0.025	1	0.067	−0.38	0.7163
1 vs 3	0.13	1	0.067	1.87	0.0935
1 vs 4	−0.30	1	0.067	−4.50	0.0015
2 vs 3	0.15	1	0.067	2.25	0.0510
2 vs 4	−0.27	1	0.067	−4.12	0.0026
3 vs 4	−0.43	1	0.067	−6.37	0.0001

Diagnostics Case Statistics

Standard Order	Actual Value	Predicted Value	Residual	Leverage	Student Residual	Cook's Distance	Outlier t
1	9.30	9.35	−0.050	0.438	−0.707	0.056	−0.686
2	9.40	9.38	0.025	0.438	0.354	0.014	0.336
3	9.60	9.67	−0.075	0.437	−1.061	0.125	−1.069
4	10.00	9.90	0.100	0.438	1.414	0.222	1.512
5	9.40	9.38	0.025	0.438	0.354	0.014	0.336
6	9.30	9.40	−0.100	0.437	−1.414	0.222	−1.512
7	9.80	9.70	0.100	0.437	1.414	0.222	1.512
8	9.90	9.93	−0.025	0.437	−0.354	0.014	−0.336
9	9.20	9.22	−0.025	0.438	−0.354	0.014	−0.336
10	9.40	9.25	0.150	0.437	2.121	0.500	2.828
11	9.50	9.55	−0.050	0.437	−0.707	0.056	−0.686
12	9.70	9.78	−0.075	0.437	−1.061	0.125	−1.069
13	9.70	9.65	0.050	0.438	0.707	0.056	0.686
14	9.60	9.68	−0.075	0.437	−1.061	0.125	−1.069
15	10.00	9.97	0.025	0.437	0.354	0.014	0.336
16	10.20	10.20	0.000	0.437	0.000	0.000	0.000

Figura 4-2 Salida de *Design-Expert* (condensada) para el ejemplo 4-1.

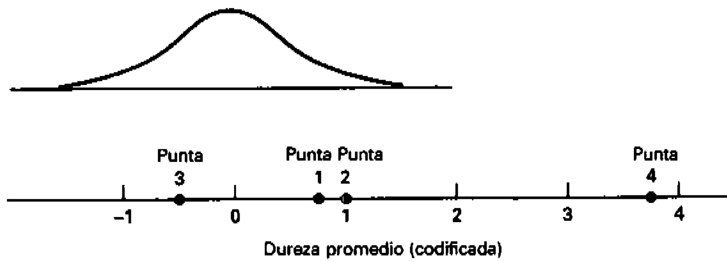


Figura 4-3 Las medias del tipo de punta en relación con una distribución t escalada con un factor de escalación $\sqrt{MS_E / b} = \sqrt{0.89 / 4} = 0.47$.

También puede usarse el procedimiento gráfico del capítulo 3 (sección 3-5.1) para comparar las medias del tipo de punta. En la figura 4-3 se grafican las cuatro medias del tipo de punta del ejemplo 4-1 en relación con una distribución t escalada con un factor de escalación $\sqrt{MS_E / b} = \sqrt{0.89 / 4} = 0.47$. Esta gráfica indica que las puntas 1, 2 y 3 producen probablemente mediciones de la dureza promedio idénticas, pero que la punta 4 produce una dureza media mucho más alta. Esta figura confirma los resultados de la prueba LSD de Fisher incluida en la salida de *Design-Expert* de la figura 4-2.

4-1.2 Verificación de la adecuación del modelo

Se ha comentado ya la importancia de verificar la adecuación del modelo supuesto. En general, deberá estarse alerta a los problemas potenciales con el supuesto de normalidad, con la desigualdad de la varianza por tratamiento o bloque, y con la interacción bloque-tratamiento. Como en el diseño completamente aleatorizado, el análisis residual es la herramienta principal que se utiliza en estos diagnósticos de verificación. En la parte inferior de la salida de *Design-Expert* de la figura 4-2 se enlistan los residuales del diseño de bloques aleatorizados. Los residuales codificados se encontrarían multiplicando estos residuales por 10. Las observaciones, los valores ajustados y los residuales de los datos codificados de la prueba de la dureza del ejemplo 4-1 son los siguientes:

y_{ij}	$\hat{y}_{ij}$	e_{ij}
−2.00	−1.50	−0.50
−1.00	−1.25	0.25
1.00	1.75	−0.75
5.00	4.00	1.00
−1.00	−1.25	0.25
−2.00	−1.00	−1.00
3.00	2.00	1.00
4.00	4.25	−0.25
−3.00	−2.75	−0.25
−1.00	−2.50	1.50
0.00	0.50	−0.50
2.00	2.75	−0.75
2.00	1.50	0.50
1.00	1.75	−0.75
5.00	4.75	0.25
7.00	7.00	0.00

En la figura 4-4 se muestra la gráfica de probabilidad normal y el diagrama de puntos de estos residuales. No hay indicios marcados de no normalidad y tampoco hay evidencia que apunte a la posible presencia de puntos atípicos. En la figura 4-5 se muestran las gráficas de los residuales por tipo de punta o

bloque 1 en conjunto es $E(y_{11}) = \mu + \tau_1 + \beta_1 = \mu + 5 + 2 = \mu + 7$. En general, el tratamiento 1 incrementa *siempre* la respuesta esperada 5 unidades sobre la suma de la media global y del efecto del bloque.

Aun cuando este modelo aditivo simple muchas veces es útil, hay situaciones en las que resulta inadecuado. Suponga, por ejemplo, que se están comparando cuatro formulaciones de un producto químico utilizando seis lotes de materia prima; los lotes de materia prima se consideran bloques. Si una impureza en el lote 2 afecta de manera adversa la formulación 2, dando como resultado un rendimiento inusualmente bajo, pero no afecta las demás formulaciones, ha ocurrido una **interacción** entre las formulaciones (o tratamientos) y los lotes (o bloques). De manera similar, pueden ocurrir interacciones entre los tratamientos y los bloques cuando la respuesta se mide en la escala incorrecta. Por lo tanto, una relación que es multiplicativa en las unidades originales, por ejemplo

$$E(y_{ij}) = \mu\tau_i\beta_j$$

es lineal o aditiva en una escala logarítmica, ya que, por ejemplo,

$$\ln E(y_{ij}) = \ln\mu + \ln\tau_i + \ln\beta_j$$

o

$$E(y_{ij}^*) = \mu^* + \tau_i^* + \beta_j^*$$

Aun cuando este tipo de interacción puede eliminarse con una transformación, no todas las interacciones pueden tratarse con tanta facilidad. Por ejemplo, una transformación no elimina la interacción formulación-lote que se señaló antes. El análisis residual y otros procedimientos de diagnóstico de verificación pueden ser útiles para detectar la no aditividad.

Si una interacción está presente, puede afectar seriamente el análisis de varianza y posiblemente lo invalide. En general, la presencia de una interacción infla el cuadrado medio del error y puede afectar adversamente la comparación de las medias de los tratamientos. En situaciones en las que ambos factores, así como su posible interacción, son de interés, deben usarse **diseños factoriales.** Estos diseños se analizan en detalle en los capítulos 5 al 9.

Tratamientos y bloques aleatorios

Aun cuando el procedimiento de prueba se ha descrito considerando los tratamientos y los bloques como factores fijos, se utiliza el mismo procedimiento de análisis si los tratamientos o los bloques (o ambos) son aleatorios. Sin embargo, hay algunas modificaciones en la interpretación de los resultados. Por ejemplo, si los bloques son aleatorios, como es con mucha frecuencia el caso, se espera que las comparaciones entre los tratamientos sean las mismas a lo largo de la **población de bloques** de la cual se seleccionaron aleatoriamente para realizar el experimento. Están también las modificaciones correspondientes en los cuadrados medios esperados. Por ejemplo, si los bloques son variables aleatorias independientes con varianza común, entonces $E(MS_{\text{Bloques}}) = \sigma^2 + a\sigma_\beta^2$, donde σ_β^2 es el componente de la varianza de los efectos de los bloques. En cualquier situación, $E(MS_{\text{Tratamientos}})$ siempre está libre de cualquier efecto de bloque, y el estadístico de prueba para la variabilidad entre los tratamientos siempre es $F_0 = MS_{\text{Tratamientos}}/MS_E$.

En situaciones en las que los bloques son aleatorios, si está presente una interacción tratamiento-bloque, las pruebas para las medias de los tratamientos no están afectadas por la interacción. La razón de ello es que los cuadrados medios esperados de los tratamientos y del error contienen *ambos* el efecto de la interacción; por consiguiente, la prueba de las diferencias en las medias de los tratamientos puede realizarse como de costumbre comparando el cuadrado medio de los tratamientos con el cuadrado medio del error. Este procedimiento no proporciona ninguna información acerca de la interacción.

Elección del tamaño de la muestra

La elección del **tamaño de la muestra,** o **número de bloques** que deben correrse, es una decisión importante cuando se usa un RCBD. Al incrementar el número de bloques, se incrementa el número de réplicas y el número de grados de libertad del error, con lo cual se aumenta la sensibilidad del diseño. Cualquiera de las técnicas descritas en el capítulo 3 (sección 3-7) para seleccionar el número de réplicas que deben correrse en un experimento completamente aleatorizado con un solo factor puede aplicarse de forma directa al RCBD. Para el caso de un factor fijo, las curvas de operación característica de la parte V del apéndice pueden usarse con

$$\Phi^2 = \frac{b\sum_{i=1}^{a}\tau_i^2}{a\sigma^2} \tag{4-14}$$

donde hay $a-1$ grados de libertad en el numerador y $(a-1)(b-1)$ grados de libertad en el denominador.

EJEMPLO 4-2 ··

Considere el problema de la prueba de la dureza del ejemplo 4-1. Suponga que quiere determinarse el número apropiado de bloques que deben correrse si el interés se encuentra en detectar una diferencia máxima real en las lecturas de la dureza media de 0.4 con una alta probabilidad, y una estimación razonable de la desviación estándar de los errores es $\sigma = 0.1$. (Estos valores se dan en las unidades *originales;* recuerde que el análisis de varianza se realizó usando datos *codificados.*) Por la ecuación 3-49, el valor mínimo de Φ^2 es (escribiendo b, el número de bloques, en lugar de n)

$$\Phi^2 = \frac{bD^2}{2a\sigma^2}$$

donde D es la diferencia máxima que quiere detectarse. Por lo tanto,

$$\Phi^2 = \frac{b(0.4)^2}{2(4)(0.1)^2} = 2.0b$$

Si se usan $b = 3$ bloques, entonces $\Phi = \sqrt{2.0b} = \sqrt{2.0(3)} = 2.45$, y hay $(a-1)(b-1) = 3(2) = 6$ grados de libertad del error. La parte V del apéndice con $v_1 = a-1 = 3$ y $\alpha = 0.05$ indica que el riesgo β de este diseño es aproximadamente 0.10 (potencia $= 1-\beta = 0.90$). Si se usan $b = 4$ bloques, $\Phi = \sqrt{2.0b} = \sqrt{2.0(4)} = 2.83$, con $(a-1)(b-1) = 3(3) = 9$ grados de libertad del error, y el riesgo β correspondiente es aproximadamente 0.03 (potencia $= 1-\beta = 0.97$). Tres o cuatro bloques darán como resultado un diseño con una alta probabilidad de detectar la diferencia entre las lecturas de la dureza media consideradas importantes. Debido a que los ejemplares de prueba (bloques) son baratos y están en disponibilidad y el costo de la prueba es bajo, el experimentador decide usar cuatro bloques.

··

Estimación de valores faltantes

Cuando se usa el RCBD, en ocasiones falta una observación en uno de los bloques. Esto puede ocurrir debido a descuido o error o por razones fuera del control del experimentador, tal como un daño inevitable a una unidad experimental. Una observación faltante introduce un nuevo problema en el análisis debido a que los tratamientos dejan de ser **ortogonales a los bloques;** es decir, no ocurren todos los tratamientos en

Tabla 4-7 Diseño de bloques completos aleatorizados para el experimento de la prueba de dureza con un valor faltante

	Ejemplar de prueba (bloque)			
Tipo de punta	1	2	3	4
1	–2	–1	1	5
2	–1	–2	x	4
3	–3	–1	0	2
4	2	1	5	7

cada uno de los bloques. Existen dos enfoques generales para el problema de los valores faltantes. El primero es un **análisis aproximado,** en el cual la observación faltante se estima y se lleva a cabo el análisis de varianza usual como si la observación estimada fuera un dato real, con los grados de libertad del error reducidos en 1. Este análisis aproximado es materia de esta sección. El segundo es un **análisis exacto,** el cual se revisa en la sección 4-1.4.

Suponga que falta la observación y_{ij} del tratamiento i en el bloque j. La observación faltante se denota como x. Como una ilustración, suponga que en el experimento de la prueba de dureza del ejemplo 4-1 el ejemplar de prueba 3 se rompió mientras se probaba la punta 2 y que no pudo obtenerse el dato para esa punta. Los datos aparecerían como en la tabla 4-7.

En general, se hará que $y'_{..}$ represente el gran total con una observación faltante, que $y'_{i.}$ represente el total del tratamiento con una observación faltante, y que $y'_{.j}$ sea el total del bloque con una observación faltante. Suponga que quiere estimarse la observación faltante x de tal modo que x tenga una participación mínima en la suma de cuadrados del error. Puesto que $SS_E = \Sigma_{i=1}^{a}\Sigma_{j=1}^{b}(y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})^2$, esto es equivalente a elegir x para minimizar

$$SS_E = \sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij}^2 - \frac{1}{b}\sum_{i=1}^{a}\left(\sum_{j=1}^{b} y_{ij}\right)^2 - \frac{1}{a}\sum_{j=1}^{b}\left(\sum_{i=1}^{a} y_{ij}\right)^2 + \frac{1}{ab}\left(\sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij}\right)^2$$

o

$$SS_E = x^2 - \frac{1}{b}(y'_{i.} + x)^2 - \frac{1}{a}(y'_{.j} + x)^2 + \frac{1}{ab}(y'_{..} + x)^2 + R \qquad (4\text{-}15)$$

donde R incluye todos los términos en los que no interviene x. A partir de $dSS_E/dx = 0$, se obtiene

$$x = \frac{ay'_{i.} + by'_{.j} - y'_{..}}{(a-1)(b-1)} \qquad (4\text{-}16)$$

como la estimación de la observación faltante.

Para los datos de la tabla 4-7, se encuentra que $y'_{2.} = 1$, $y'_{.3} = 6$ y $y'_{..} = 17$. Por lo tanto, por la ecuación 4-16,

$$x \equiv y_{23} = \frac{4(1) + 4(6) - 17}{(3)(3)} = 1.22$$

Ahora puede realizarse el análisis de varianza común utilizando $y_{23} = 1.22$ y reduciendo los grados de libertad del error en 1. El análisis de varianza se muestra en la tabla 4-8. Compare los resultados de este análisis aproximado con los resultados obtenidos para el conjunto de datos completo (tabla 4-5).

Tabla 4-8 Análisis de varianza aproximado del ejemplo 4-1 con un valor faltante

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0	Valor P
Tipo de punta	39.98	3	13.33	17.12	0.0008
Ejemplares de prueba (bloques)	79.53	3	26.51		
Error	6.22	8	0.78		
Total	125.73	14			

Si son varias las observaciones faltantes, pueden estimarse escribiendo la suma de cuadrados del error como una función de los valores faltantes, derivando con respecto a cada valor faltante, igualando los resultados con cero y resolviendo las ecuaciones resultantes. De manera alternativa, puede usarse la ecuación 4-16 de manera iterativa para estimar los valores faltantes. Para ilustrar el enfoque iterativo, suponga que faltan dos valores. Se estima arbitrariamente el primer valor faltante y después se usa este valor junto con los datos reales y la ecuación 4-16 para estimar el segundo. Entonces puede usarse la ecuación 4-16 para volver a estimar el primer valor faltante, y después de esto, puede volver a estimarse el segundo. Este proceso se continúa hasta que se obtiene la convergencia. En cualquier problema con valores faltantes, los grados de libertad del error se reducen en una unidad por cada observación faltante.

4-1.4 Estimación de los parámetros del modelo y la prueba general de significación de la regresión

Si tanto los tratamientos como los bloques son fijos, los parámetros del RCBD pueden estimarse por mínimos cuadrados. Recuerde que el modelo estadístico lineal es

$$y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij} \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, b \end{cases} \tag{4-17}$$

Al aplicar las reglas de la sección 3-9.2 para encontrar las ecuaciones normales del modelo de un diseño experimental, se obtiene

$$\begin{array}{lllllllllll}
\mu: & ab\hat{\mu} & + b\hat{\tau}_1 & + b\hat{\tau}_2 & + \cdots & + b\hat{\tau}_a & + a\hat{\beta}_1 & + a\hat{\beta}_2 & + \cdots + & a\hat{\beta}_b & = y_{..} \\
\tau_1: & b\hat{\mu} & + b\hat{\tau}_1 & & & & + \hat{\beta}_1 & + \hat{\beta}_2 & + \cdots + & \hat{\beta}_b & = y_{1.} \\
\tau_2: & b\hat{\mu} & & + b\hat{\tau}_2 & & & + \hat{\beta}_1 & + \hat{\beta}_2 & + \cdots + & \hat{\beta}_b & = y_{2.} \\
\vdots & & & & & & \vdots & & & & \vdots \\
\tau_a: & b\hat{\mu} & & & & + b\hat{\tau}_a & + \hat{\beta}_1 & + \hat{\beta}_2 & + \cdots + & \hat{\beta}_b & = y_{a.} \\
\beta_1: & a\hat{\mu} & + \hat{\tau}_1 & + \hat{\tau}_2 & + \cdots & + \hat{\tau}_a & + a\hat{\beta}_1 & & & & = y_{.1} \\
\beta_2: & a\hat{\mu} & + \hat{\tau}_1 & + \hat{\tau}_2 & + \cdots & + \hat{\tau}_a & & + a\hat{\beta}_2 & & & = y_{.2} \\
\vdots & & & & & & \vdots & & & & \vdots \\
\beta_b: & a\hat{\mu} & + \hat{\tau}_1 & + \hat{\tau}_2 & + \cdots & + \hat{\tau}_a & & & + & a\hat{\beta}_b & = y_{.b}
\end{array} \tag{4-18}$$

Observe que la suma de la segunda a la $(a + 1)$-ésima ecuaciones de la ecuación 4-18 es la primera ecuación normal, como también es el caso de las b últimas ecuaciones. Por lo tanto, hay dos dependencias

lineales en las ecuaciones normales, lo cual implica que deben imponerse dos restricciones para resolver la ecuación 4-18. Las restricciones usuales son

$$\sum_{i=1}^{a} \hat{\tau}_i = 0 \qquad \sum_{j=1}^{b} \hat{\beta}_j = 0 \tag{4-19}$$

Al utilizar estas restricciones, las ecuaciones normales se simplifican considerablemente. De hecho, quedan como

$$\begin{aligned} ab\hat{\mu} &= y_{..} \\ b\hat{\mu}+b\hat{\tau}_i &= y_{i.} \qquad i=1,2,\ldots,a \\ a\hat{\mu}+a\hat{\beta}_j &= y_{.j} \qquad j=1,2,\ldots,b \end{aligned} \tag{4-20}$$

cuya solución es

$$\begin{aligned} \hat{\mu} &= \bar{y}_{..} \\ \hat{\tau}_i &= \bar{y}_{i.} - \bar{y}_{..} \qquad i=1,2,\ldots,a \\ \hat{\beta}_j &= \bar{y}_{.j} - \bar{y}_{..} \qquad j=1,2,\ldots,b \end{aligned} \tag{4-21}$$

Al utilizar la solución de la ecuación normal de la ecuación 4-21, puede encontrarse el valor estimado o ajustado de y_{ij} como

$$\begin{aligned} \hat{y}_{ij} &= \hat{\mu}+\hat{\tau}_i+\hat{\beta}_j \\ &= \bar{y}_{..}+(\bar{y}_{i.}-\bar{y}_{..})+(\bar{y}_{.j}-\bar{y}_{..}) \\ &= \bar{y}_{i.}+\bar{y}_{.j}-\bar{y}_{..} \end{aligned}$$

Este resultado se usó anteriormente en la ecuación 4-13 para calcular los residuales de un diseño de bloques aleatorizados.

La prueba general de significación de la regresión puede usarse para desarrollar el análisis de varianza del diseño de bloques completos aleatorizados. Al utilizar la solución de las ecuaciones normales dada por la ecuación 4-21, la reducción en la suma de cuadrados para ajustar el modelo completo es

$$\begin{aligned} R(\mu,\tau,\beta) &= \hat{\mu}y_{..}+\sum_{i=1}^{a}\hat{\tau}_i y_{i.}+\sum_{j=1}^{b}\hat{\beta}_j y_{.j} \\ &= \bar{y}_{..}y_{..}+\sum_{i=1}^{a}(\bar{y}_{i.}-\bar{y}_{..})y_{i.}+\sum_{j=1}^{b}(\bar{y}_{.j}-\bar{y}_{..})y_{.j} \\ &= \frac{y_{..}^2}{ab}+\sum_{i=1}^{a}\bar{y}_{i.}y_{i.}-\frac{y_{..}^2}{ab}+\sum_{j=1}^{b}\bar{y}_{.j}y_{.j}-\frac{y_{..}^2}{ab} \\ &= \sum_{i=1}^{a}\frac{y_{i.}^2}{b}+\sum_{j=1}^{b}\frac{y_{.j}^2}{a}-\frac{y_{..}^2}{ab} \end{aligned}$$

con $a + b - 1$ grados de libertad, y la suma de cuadrados del error es

$$\begin{aligned} SS_E &= \sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij}^2 - R(\mu,\tau,\beta) \\ &= \sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij}^2 - \sum_{i=1}^{a}\frac{y_{i.}^2}{b} - \sum_{j=1}^{b}\frac{y_{.j}^2}{a}+\frac{y_{..}^2}{ab} \\ &= \sum_{i=1}^{a}\sum_{j=1}^{b}(y_{ij}-\bar{y}_{i.}-\bar{y}_{.j}+\bar{y}_{..})^2 \end{aligned}$$

con $(a-1)(b-1)$ grados de libertad. Compare esta última ecuación con SS_E en la ecuación 4-7.

Para probar la hipótesis $H_0{:}\tau_i = 0$, el modelo reducido es

$$y_{ij} = \mu + \beta_j + \varepsilon_{ij}$$

que es un análisis de varianza de un solo factor. Por analogía con la ecuación 3-5, la reducción en la suma de cuadrados para ajustar el modelo reducido es

$$R(\mu, \beta) = \sum_{j=1}^{b} \frac{y_{.j}^2}{a}$$

que tiene b grados de libertad. Por lo tanto, la suma de cuadrados debida a $\{\tau_i\}$ después de ajustar μ y $\{\beta_j\}$ es

$$\begin{aligned} R(\tau|\mu, \beta) &= R(\mu, \tau, \beta) - R(\mu, \beta) \\ &= R(\text{Modelo completo}) - R(\text{Modelo reducido}) \\ &= \sum_{i=1}^{a} \frac{y_{i.}^2}{b} + \sum_{j=1}^{b} \frac{y_{.j}^2}{a} - \frac{y_{..}^2}{ab} - \sum_{j=1}^{b} \frac{y_{.j}^2}{a} \\ &= \sum_{i=1}^{a} \frac{y_{i.}^2}{b} - \frac{y_{..}^2}{ab} \end{aligned}$$

expresión que se identifica como la suma de cuadrados de los tratamientos con $a-1$ grados de libertad (ecuación 4-10).

La suma de cuadrados de los bloques se obtiene ajustando el modelo reducido

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij}$$

que también es un análisis de un solo factor. De nueva cuenta, por analogía con la ecuación 3-5, la reducción en la suma de cuadrados para ajustar este modelo es

$$R(\mu, \tau) = \sum_{i=1}^{a} \frac{y_{i.}^2}{b}$$

con a grados de libertad. La suma de cuadrados de los bloques $\{\beta_j\}$ después de ajustar μ y $\{\tau_i\}$ es

$$\begin{aligned} R(\beta|\mu, \tau) &= R(\mu, \tau, \beta) - R(\mu, \tau) \\ &= \sum_{i=1}^{a} \frac{y_{i.}^2}{b} + \sum_{j=1}^{b} \frac{y_{.j}^2}{a} - \frac{y_{..}^2}{ab} - \sum_{i=1}^{a} \frac{y_{i.}^2}{b} \\ &= \sum_{j=1}^{b} \frac{y_{.j}^2}{a} - \frac{y_{..}^2}{ab} \end{aligned}$$

con $b-1$ grados de libertad, la cual se había dado anteriormente como la ecuación 4-11.

Se han desarrollado las sumas de cuadrados de los tratamientos, de los bloques y del error en el diseño de bloques completos aleatorizados utilizando la prueba general de significación de la regresión. Aun cuando la prueba general de significación de la regresión no se usaría ordinariamente para hacer el análisis real de los datos en un bloque completo aleatorizado, en ocasiones el procedimiento resulta útil en diseños de bloques aleatorizados más generales, como los que se revisan en la sección 4-4.

Análisis exacto del problema del valor faltante

En la sección 4-1.3 se presentó un procedimiento aproximado para solucionar las observaciones faltantes en el RCBD. Este análisis aproximado consiste en estimar el valor faltante de tal modo que se minimice el cuadrado medio del error. Puede demostrarse que el análisis aproximado produce un cuadrado medio sesgado de los tratamientos en el sentido de que $E(MS_{\text{Tratamientos}})$ es mayor que $E(MS_E)$ si la hipótesis nula es verdadera. Por consiguiente, se reportan demasiados resultados significativos.

El problema del valor faltante puede analizarse exactamente utilizando la prueba general de significación de la regresión. El valor faltante hace que el diseño sea **no balanceado,** y dado que no todos los tratamientos ocurren en todos los bloques, se dice que los tratamientos y los bloques no son **ortogonales.** Este método de análisis también se usa en tipos más generales de diseños de bloques aleatorizados; el tema se analiza con mayor amplitud en la sección 4-4. En el problema 4-26 se le pide al lector que realice el análisis exacto para un diseño de bloques completos aleatorizados con un valor faltante.

4-2 DISEÑO DE CUADRADO LATINO

En la sección 4-1 se introdujo el diseño de bloques completos aleatorizados como un diseño para reducir el error residual de un experimento al eliminar la variabilidad debida a una variable perturbadora conocida y controlable. Hay otros tipos de diseños que utilizan el principio de la formación de bloques. Por ejemplo, suponga que un experimentador estudia los efectos que tienen cinco formulaciones diferentes de la carga propulsora utilizada en los sistemas de expulsión de la tripulación de un avión basado en la rapidez de combustión. Cada formulación se hace con un lote de materia prima que sólo alcanza para probar cinco formulaciones. Además, las formulaciones son preparadas por varios operadores, y puede haber diferencias sustanciales en las habilidades y experiencia de los operadores. Por lo tanto, al parecer hay dos factores perturbadores que serán "calculados en promedio" en el diseño: los lotes de materia prima y los operadores. El diseño apropiado para este problema consiste en probar cada formulación exactamente una vez con cada uno de los cinco operadores. Al diseño resultante, ilustrado en la tabla 4-9, se le llama **diseño de cuadrado latino.** Observe que el diseño es un arreglo cuadrado y que las cinco formulaciones (o tratamientos) se denotan por las letras latinas *A, B, C, D* y *E*; de ahí el nombre de cuadrado latino. Se observa que tanto los lotes de materia prima (renglones) como los operadores (columnas) son ortogonales a los tratamientos.

El diseño de cuadrado latino se usa para eliminar dos fuentes de variabilidad perturbadora; es decir, permite hacer la formación de bloques sistemática en dos direcciones. Por lo tanto, los renglones y las columnas representan en realidad *dos restricciones sobre la aleatorización*. En general, un cuadrado latino para p factores, o cuadrado latino $p \times p$, es un cuadrado con p renglones y p columnas. Cada una de las p^2

Tabla 4-9 Diseño del cuadrado latino para el problema de la carga propulsora

Lotes de materia prima	Operadores				
	1	2	3	4	5
1	$A = 24$	$B = 20$	$C = 19$	$D = 24$	$E = 24$
2	$B = 17$	$C = 24$	$D = 30$	$E = 27$	$A = 36$
3	$C = 18$	$D = 38$	$E = 26$	$A = 27$	$B = 21$
4	$D = 26$	$E = 31$	$A = 26$	$B = 23$	$C = 22$
5	$E = 22$	$A = 30$	$B = 20$	$C = 29$	$D = 31$

celdas resultantes contiene una de las p letras que corresponde a los tratamientos, y cada letra ocurre una y sólo una vez en cada renglón y columna. Algunos ejemplos de cuadrados latinos son

4 × 4	5 × 5	6 × 6
$A\,B\,D\,C$	$A\,D\,B\,E\,C$	$A\,D\,C\,E\,B\,F$
$B\,C\,A\,D$	$D\,A\,C\,B\,E$	$B\,A\,E\,C\,F\,D$
$C\,D\,B\,A$	$C\,B\,E\,D\,A$	$C\,E\,D\,F\,A\,B$
$D\,A\,C\,B$	$B\,E\,A\,C\,D$	$D\,C\,F\,B\,E\,A$
	$E\,C\,D\,A\,B$	$F\,B\,A\,D\,C\,E$
		$E\,F\,B\,A\,D\,C$

El **modelo estadístico** de un cuadrado latino es

$$y_{ijk} = \mu + \alpha_i + \tau_j + \beta_k + \varepsilon_{ijk} \begin{cases} i = 1, 2, \ldots, p \\ j = 1, 2, \ldots, p \\ k = 1, 2, \ldots, p \end{cases} \tag{4-22}$$

donde y_{ijk} es la observación en el renglón i-ésimo y la columna k-ésima para el tratamiento j-ésimo, μ es la media global, α_i es el efecto del renglón i-ésimo, τ_j es el efecto del tratamiento j-ésimo, β_k es el efecto de la columna k-ésima, y ε_{ijk} es el error aleatorio. Observe que se trata de un **modelo de los efectos.** El modelo es completamente **aditivo;** es decir, no hay interacción entre renglones, columnas y tratamientos. Puesto que hay una sola observación en cada celda, sólo se necesitan dos de los tres subíndices i, j y k para denotar una observación particular. Por ejemplo, con referencia al problema de la carga propulsora de la tabla 4-9, si $i = 2$ y $k = 3$, se encuentra automáticamente que $j = 4$ (formulación D), y si $i = 1$ y $j = 3$ (formulación C), se encuentra que $k = 3$. Ésta es una consecuencia de que cada tratamiento aparezca una vez exactamente en cada renglón y columna.

El análisis de varianza consiste en hacer la partición de la suma de cuadrados total de las $N = p^2$ observaciones en los componentes de los renglones, las columnas, los tratamientos y el error, por ejemplo,

$$SS_T = SS_{\text{Renglones}} + SS_{\text{Columnas}} + SS_{\text{Tratamientos}} + SS_E \tag{4-23}$$

con los respectivos grados de libertad

$$p^2 - 1 = p - 1 + p - 1 + p - 1 + (p-2)(p-1)$$

Bajo el supuesto usual de que ε_{ijk} es NID(0, σ^2), cada suma de cuadrados del lado derecho de la ecuación 4-23 es, al dividir por σ^2, una variable aleatoria ji-cuadrada con una distribución independiente. El estadístico apropiado para probar que no hay diferencias en las medias de los tratamientos es

$$F_0 = \frac{MS_{\text{Tratamientos}}}{MS_E}$$

que se distribuye como $F_{p-1,(p-2)(p-1)}$ bajo la hipótesis nula. También puede probarse la ausencia de efectos de los renglones o la ausencia de efectos de las columnas formando el cociente de $MS_{\text{Renglones}}$ o MS_{Columnas} con MS_E. Sin embargo, puesto que los renglones y las columnas representan restricciones sobre la aleatorización, estas pruebas quizá no sean apropiadas.

En la tabla 4-10 se presenta el procedimiento de cálculo para el análisis de varianza. Por las fórmulas de cálculo para las sumas de cuadrados, se observa que el análisis es una extensión simple del RCBD, con la suma de cuadrados resultante de los renglones obtenida a partir de los totales de los renglones.

Tabla 4-10 Análisis de varianza del diseño del cuadrado latino

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0
Tratamientos	$SS_{\text{Tratamientos}} = \frac{1}{p}\sum_{j=1}^{p} y_{.j.}^2 - \frac{y_{...}^2}{N}$	$p-1$	$\frac{SS_{\text{Tratamientos}}}{p-1}$	$F_0 = \frac{MS_{\text{Tratamientos}}}{MS_E}$
Renglones	$SS_{\text{Renglones}} = \frac{1}{p}\sum_{i=1}^{p} y_{i..}^2 - \frac{y_{...}^2}{N}$	$p-1$	$\frac{SS_{\text{Renglones}}}{p-1}$	
Columnas	$SS_{\text{Columnas}} = \frac{1}{p}\sum_{k=1}^{p} y_{..k}^2 - \frac{y_{...}^2}{N}$	$p-1$	$\frac{SS_{\text{Columnas}}}{p-1}$	
Error	SS_E (por sustracción)	$(p-2)(p-1)$	$\frac{SS_E}{(p-2)(p-1)}$	
Total	$SS_T = \sum_i \sum_j \sum_k y_{ijk}^2 - \frac{y_{...}^2}{N}$	p^2-1		

EJEMPLO 4-3

Considere el problema de la carga propulsora descrito previamente, donde tanto los lotes de materia prima como los operadores representan restricciones sobre la aleatorización. El diseño para este experimento, el cual se muestra en la tabla 4-9, es un cuadrado latino 5 × 5. Después de codificar los datos restando 25 de cada observación, se obtienen los datos de la tabla 4-11. Las sumas de cuadrados del total de los lotes (renglones) y los operadores (columnas) se calculan de la siguiente manera:

$$
\begin{aligned}
SS_T &= \sum_i \sum_j \sum_k y_{ijk}^2 - \frac{y_{...}^2}{N} \\
&= 680 - \frac{(10)^2}{25} = 676.00 \\
SS_{\text{Lotes}} &= \frac{1}{p}\sum_{i=1}^{p} y_{i..}^2 - \frac{y_{...}^2}{N} \\
&= \frac{1}{5}[(-14)^2 + 9^2 + 5^2 + 3^2 + 7^2] - \frac{(10)^2}{25} = 68.00 \\
SS_{\text{Operadores}} &= \frac{1}{p}\sum_{k=1}^{p} y_{..k}^2 - \frac{y_{...}^2}{N} \\
&= \frac{1}{5}[(-18)^2 + 18^2 + (-4)^2 + 5^2 + 9^2] - \frac{(10)^2}{25} = 150.00
\end{aligned}
$$

Tabla 4-11 Datos codificados para el problema de la carga propulsora

Lotes de materia prima	Operadores 1	2	3	4	5	$y_{i..}$
1	$A = -1$	$B = -5$	$C = -6$	$D = -1$	$E = -1$	−14
2	$B = -8$	$C = -1$	$D = 5$	$E = 2$	$A = 11$	9
3	$C = -7$	$D = 13$	$E = 1$	$A = 2$	$B = -4$	5
4	$D = 1$	$E = 6$	$A = 1$	$B = -2$	$C = -3$	3
5	$E = -3$	$A = 5$	$B = -5$	$C = 4$	$D = 6$	7
$y_{..k}$	−18	18	−4	5	9	$10 = y_{...}$

Tabla 4-12 Análisis de varianza del experimento de la carga propulsora

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0	Valor P
Formulaciones	330.00	4	82.50	7.73	0.0025
Lotes de materia prima	68.00	4	17.00		
Operadores	150.00	4	37.50		
Error	128.00	12	10.67		
Total	676.00	24			

Los totales para los tratamientos (las letras latinas) son

Letra latina	**Total del tratamiento**
A	$y_{.1.} = 18$
B	$y_{.2.} = -24$
C	$y_{.3.} = -13$
D	$y_{.4.} = 24$
E	$y_{.5.} = 5$

La suma de cuadrados que resulta de las formulaciones se calcula a partir de estos totales como

$$SS_{\text{Formulaciones}} = \frac{1}{p}\sum_{j=1}^{p} y_{.j.}^2 - \frac{y_{...}^2}{N}$$

$$= \frac{18^2 + (-24)^2 + (-13)^2 + 24^2 + 5^2}{5} - \frac{(10)^2}{25} = 330.00$$

La suma de cuadrados del error se encuentra por sustracción:

$$SS_E = SS_T - SS_{\text{Lotes}} - SS_{\text{Operadores}} - SS_{\text{Formulaciones}}$$

$$= 676.00 - 68.00 - 150.00 - 330.00 = 128.00$$

El análisis de varianza se resume en la tabla 4-12. Se concluye que hay una diferencia significativa en la rapidez de combustión media generada por las diferentes formulaciones de la carga propulsora. También hay indicios de que hay diferencias entre los operadores, por lo que la formación de bloques de este factor fue una buena precaución. No hay evidencia sólida de una diferencia entre los lotes de materia prima, por lo que al parecer en este experimento particular hubo una preocupación innecesaria en esta fuente de variabilidad. Sin embargo, la formación de bloques de los lotes de materia prima es por lo general una buena idea.

Como en cualquier problema de diseño, el experimentador debería investigar la adecuación del modelo inspeccionando y graficando los residuales. Para un cuadrado latino, los residuales están dados por

$$e_{ijk} = y_{ijk} - \hat{y}_{ijk}$$

$$= y_{ijk} - \bar{y}_{i..} - \bar{y}_{.j.} - \bar{y}_{..k} + 2\bar{y}_{...}$$

El lector deberá encontrar los residuales del ejemplo 4-4 y construir las gráficas apropiadas.

Tabla 4-13 Cuadrados latinos estándares y número de cuadrados latinos de varios tamaños[a]

Tamaño	3 × 3	4 × 4	5 × 5	6 × 6	7 × 7	p × p
Ejemplos de cuadrados estándares	*ABC* *BCA* *CAB*	*ABCD* *BCDA* *CDAB* *DABC*	*ABCDE* *BAECD* *CDAEB* *DEBAC* *ECDBA*	*ABCDEF* *BCFADE* *CFBEAD* *DEABFC* *EADFCB* *FDECBA*	*ABCDEFG* *BCDEFGA* *CDEFGAB* *DEFGABC* *EFGABCD* *FGABCDE* *GABCDEF*	*ABC … P* *BCD … A* *CDE … B* ⋮ *PAB* … (*P*–1)
Número de cuadrados estándares	1	4	56	9408	16,942,080	–
Número total de cuadrados latinos	12	576	161,280	818,851,200	61,479,419,904,000	$p!(p-1)!$ × (número de cuadrados estándares)

[a]Parte de la información de esta tabla se encuentra en *Statistical Tables for Biological, Agricultural and Medical Research*, 4a. edición, de R.A. Fisher y F. Yates, Oliver & Boyd, Edimburgo. Es poco lo que se sabe de las propiedades de los cuadrados latinos más grandes que 7 × 7.

A un cuadrado latino en el que el primer renglón y la primera columna constan de letras escritas en orden alfabético se le llama **cuadrado latino estándar,** que es el diseño que se utilizó en el ejemplo 4-3. Siempre es posible obtener un cuadrado latino estándar escribiendo el primer renglón en orden alfabético y escribiendo después cada renglón sucesivo como la sucesión de letras que están justo arriba, recorridas un lugar a la izquierda. En la tabla 4-13 se resumen varios hechos importantes acerca de los cuadrados latinos y de los cuadrados latinos estándares.

Como con cualquier diseño experimental, las observaciones del cuadrado latino deberán tomarse de manera aleatoria. El procedimiento de aleatorización correcto es seleccionando al azar el cuadrado empleado. Como se observa en la tabla 4-13, hay un gran número de cuadrados latinos de un tamaño particular, por lo que es imposible enumerar todos los cuadrados y seleccionar uno al azar. El procedimiento usual es seleccionar un cuadrado latino de una tabla de estos diseños, como en Fisher y Yates [45], y después arreglar al azar el orden de los renglones, las columnas y las letras. Esto se analiza con mayor detalle en Fisher y Yates [45].

Ocasionalmente, falta una observación en un cuadrado latino. Para un cuadrado latino $p \times p$, el valor faltante puede estimarse con

$$y_{ijk} = \frac{p(y'_{i..} + y'_{.j.} + y'_{..k}) - 2y'_{...}}{(p-2)(p-1)} \tag{4-24}$$

donde las primas indican los totales del renglón, la columna y el tratamiento con el valor faltante, y $y'_{...}$ es el gran total con el valor faltante.

Los cuadrados latinos pueden ser útiles en situaciones en las que los renglones y las columnas representan los factores que el experimentador en realidad quiere estudiar y en las que no hay restricciones sobre la aleatorización. Por lo tanto, los tres factores (renglones, columnas y letras), cada uno con p niveles, pueden investigarse en sólo p^2 corridas. En este diseño se supone que no existe interacción entre los factores. Se abundará más adelante sobre el tema de la interacción.

Réplicas de cuadrados latinos

Una desventaja de los cuadrados latinos pequeños es que proporcionan un número relativamente pequeño de grados de libertad del error. Por ejemplo, un cuadrado latino 3 × 3 sólo tiene dos grados de libertad del error, un cuadrado latino 4 × 4 sólo tiene seis grados de libertad del error, etc. Cuando se usan cuadrados latinos pequeños, con frecuencia es deseable hacer réplicas de los mismos para incrementar los grados de libertad del error.

Tabla 4-14 Análisis de varianza de un cuadrado latino con réplicas, caso 1

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0
Tratamientos	$\frac{1}{np}\sum_{j=1}^{p} y_{.j..}^2 - \frac{y_{....}^2}{N}$	$p-1$	$\frac{SS_{\text{Tratamientos}}}{p-1}$	$\frac{MS_{\text{Tratamientos}}}{MS_E}$
Renglones	$\frac{1}{np}\sum_{i=1}^{p} y_{i...}^2 - \frac{y_{....}^2}{N}$	$p-1$	$\frac{SS_{\text{Renglones}}}{p-1}$	
Columnas	$\frac{1}{np}\sum_{k=1}^{p} y_{..k.}^2 - \frac{y_{....}^2}{N}$	$p-1$	$\frac{SS_{\text{Columnas}}}{p-1}$	
Réplicas	$\frac{1}{p^2}\sum_{l=1}^{n} y_{...l}^2 - \frac{y_{....}^2}{N}$	$n-1$	$\frac{SS_{\text{Réplicas}}}{n-1}$	
Error	Sustracción	$(p-1)[n(p+1)-3]$	$\frac{SS_E}{(p-1)[n(p+1)-3]}$	
Total	$\sum\sum\sum\sum y_{ijkl}^2 - \frac{y_{....}^2}{N}$	np^2-1		

Existen varias maneras de hacer réplicas de un cuadrado latino. Para ilustrar este punto, suponga que se hacen n réplicas del cuadrado latino 5×5 utilizado en el ejemplo 4-3. Esto podría haberse hecho de la manera siguiente:

1. Usando los mismos lotes y operadores en cada réplica.
2. Usando los mismos lotes pero operadores diferentes en cada réplica (o, de manera equivalente, usando los mismos operadores pero lotes diferentes).
3. Usando diferentes lotes y diferentes operadores.

El análisis de varianza depende del método utilizado para hacer las réplicas.

Considere el caso 1, donde en cada réplica se usan los mismos niveles de los factores para la formación de bloques en los renglones y las columnas. Sea y_{ijkl} la observación del renglón i, el tratamiento j, la columna k y la réplica l. Hay en total $N = np^2$ observaciones. El análisis de varianza se resume en la tabla 4-14.

Considere ahora el caso 2 y suponga que en cada réplica se usan nuevos lotes de materia prima pero los mismos operadores. Por lo tanto, hay ahora cinco nuevos renglones (en general, p nuevos renglones)

Tabla 4-15 Análisis de varianza de un cuadrado latino con réplicas, caso 2

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0
Tratamientos	$\frac{1}{np}\sum_{j=1}^{p} y_{.j..}^2 - \frac{y_{....}^2}{N}$	$p-1$	$\frac{SS_{\text{Tratamientos}}}{p-1}$	$\frac{MS_{\text{Tratamientos}}}{MS_E}$
Renglones	$\frac{1}{p}\sum_{l=1}^{n}\sum_{i=1}^{p} y_{i..l}^2 - \sum_{l=1}^{n} \frac{y_{...l}^2}{p^2}$	$n(p-1)$	$\frac{SS_{\text{Renglones}}}{n(p-1)}$	
Columnas	$\frac{1}{np}\sum_{k=1}^{p} y_{..k.}^2 - \frac{y_{....}^2}{N}$	$p-1$	$\frac{SS_{\text{Columnas}}}{p-1}$	
Réplicas	$\frac{1}{p^2}\sum_{l=1}^{n} y_{...l}^2 - \frac{y_{....}^2}{N}$	$n-1$	$\frac{SS_{\text{Réplicas}}}{n-1}$	
Error	Sustracción	$(p-1)(np-1)$	$\frac{SS_E}{(p-1)(np-1)}$	
Total	$\sum_i\sum_j\sum_k\sum_l y_{ijkl}^2 - \frac{y_{....}^2}{N}$	np^2-1		

Tabla 4-16 Análisis de varianza de un cuadrado latino con réplicas, caso 3

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0
Tratamientos	$\frac{1}{np}\sum_{j=1}^{p} y_{.j..}^2 - \frac{y_{....}^2}{N}$	$p-1$	$\frac{SS_{\text{Tratamientos}}}{p-1}$	$\frac{MS_{\text{Tratamientos}}}{MS_E}$
Renglones	$\frac{1}{p}\sum_{l=1}^{n}\sum_{i=1}^{p} y_{i..l}^2 - \sum_{l=1}^{n} \frac{y_{...l}^2}{p^2}$	$n(p-1)$	$\frac{SS_{\text{Renglones}}}{n(p-1)}$	
Columnas	$\frac{1}{p}\sum_{l=1}^{n}\sum_{k=1}^{p} y_{..kl}^2 - \sum_{l=1}^{n} \frac{y_{...l}^2}{p^2}$	$n(p-1)$	$\frac{SS_{\text{Columnas}}}{n(p-1)}$	
Réplicas	$\frac{1}{p^2}\sum_{l=1}^{n} y_{...l}^2 - \frac{y_{....}^2}{N}$	$n-1$	$\frac{SS_{\text{Réplicas}}}{n-1}$	
Error	Sustracción	$(p-1)[n(p-1)-1]$	$\frac{SS_E}{(p-1)[n(p-1)-1]}$	
Total	$\sum_i\sum_j\sum_k\sum_l y_{ijkl}^2 - \frac{y_{....}^2}{N}$	np^2-1		

dentro de cada réplica. El análisis de varianza se resume en la tabla 4-15. Observe que la fuente de variación de los renglones mide en realidad la variación entre los renglones dentro de las n réplicas.

Por último, considere el caso 3, donde se usan nuevos lotes de materia prima y nuevos operadores en cada réplica. Ahora la variación que resulta tanto de los renglones como de las columnas mide la variación que resulta de estos factores dentro de las réplicas. El análisis de varianza se resume en la tabla 4-16.

Hay otros enfoques para analizar cuadrados latinos con réplicas que permiten la presencia de algunas interacciones entre tratamientos y cuadrados (referirse al problema 4-19).

Diseños alternados y diseños balanceados para efectos residuales

Ocasionalmente aparece un problema en el que los periodos son uno de los factores del experimento. En general, hay p tratamientos que deben probarse en p periodos utilizando np unidades experimentales. Por ejemplo, un analista del desempeño humano está estudiando el efecto de dos fluidos de restitución para la deshidratación en 20 sujetos. En el primer periodo, a la mitad de los sujetos (elegidos al azar) se le administra el fluido A y a la otra mitad el fluido B. Al término del periodo se mide la respuesta, y se deja transcurrir un lapso en el que se elimina cualquier efecto fisiológico de los fluidos. Después el experimentador hace que los sujetos que tomaron el fluido A tomen el fluido B y aquellos que tomaron el fluido B tomen el fluido A. A este diseño se le llama **diseño alternado** o **entrecruzado.** Se analiza como un conjunto de 10 cuadrados latinos con dos renglones (los periodos) y dos tratamientos (los tipos de fluido). Las dos columnas en cada uno de los 10 cuadrados corresponden a los sujetos.

En la figura 4-7 se muestra la disposición de este diseño. Observe que los renglones del cuadrado latino representan a los periodos y que las columnas representan a los sujetos. Los 10 sujetos que recibieron primero el fluido A (1, 4, 6, 7, 9, 12, 13, 15, 17 y 19) se determinaron al azar.

	Cuadrados latinos																			
	I		II		III		IV		V		VI		VII		VIII		IX		X	
Sujeto	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Periodo 1	*A*	*B*	*B*	*A*	*B*	*A*	*A*	*B*	*A*	*B*	*B*	*A*	*A*	*B*	*A*	*B*	*A*	*B*	*A*	*B*
Periodo 2	*B*	*A*	*A*	*B*	*A*	*B*	*B*	*A*	*B*	*A*	*A*	*B*	*B*	*A*	*B*	*A*	*B*	*A*	*B*	*A*

Figura 4-7 Un diseño alternado.

Tabla 4-17 Análisis de varianza del diseño alternado de la figura 4-7

Fuente de variación	Grados de libertad
Sujetos (columnas)	19
Periodos (renglones)	1
Fluidos (letras)	1
Error	18
Total	39

En la tabla 4-17 se resume un análisis de varianza. La suma de cuadrados de los sujetos se calcula como la suma de cuadrados entre los totales de los 20 sujetos corregida, la suma de cuadrados de los periodos es la suma de cuadrados entre los renglones corregida, y la suma de cuadrados de los fluidos se calcula como la suma de cuadrados entre los totales de las letras corregida. Para más detalles del análisis estadístico de estos diseños, ver Cochran y Cox [26], John [61d] y Anderson y McLean [2].

También es posible emplear diseños tipo cuadrado latino para experimentos en los que los tratamientos tienen un **efecto residual**; es decir, por ejemplo, si los datos del fluido B en el periodo 2 siguen reflejando algún efecto del fluido A tomado en el periodo 1. En Cochran y Cox [26] y John [61d] se estudian en detalle los diseños balanceados para efectos residuales.

4-3 DISEÑO DE CUADRADO GRECOLATINO

Considere un cuadrado latino $p \times p$ al cual se le superpone un segundo cuadrado latino $p \times p$ en el que los tratamientos se denotan con letras griegas. Si cuando se hace la superposición los dos cuadrados tienen la propiedad de que cada letra griega aparece una y sólo una vez con cada letra latina, se dice que los dos cuadrados latinos son *ortogonales*, y al diseño obtenido se le llama **cuadrado grecolatino.** En la tabla 4-18 se muestra un ejemplo de un cuadrado grecolatino 4×4.

El diseño de cuadrado grecolatino puede usarse para controlar sistemáticamente tres fuentes de variabilidad extraña, es decir, para hacer la formación de bloques en *tres* direcciones. El diseño permite la investigación de cuatro factores (renglones, columnas, letras latinas y letras griegas), cada una con p niveles en sólo p^2 corridas. Existen cuadrados grecolatinos para toda $p \geq 3$, excepto $p = 6$.

Tabla 4-18 Diseño del cuadrado grecolatino 4×4

	Columna			
Renglón	1	2	3	4
1	$A\alpha$	$B\beta$	$C\gamma$	$D\delta$
2	$B\delta$	$A\gamma$	$D\beta$	$C\alpha$
3	$C\beta$	$D\alpha$	$A\delta$	$B\gamma$
4	$D\gamma$	$C\delta$	$B\alpha$	$A\beta$

Tabla 4-19 Análisis de varianza de un diseño del cuadrado grecolatino

Fuente de variación	Suma de cuadrados	Grados de libertad
Tratamientos con letras latinas	$SS_L = \frac{1}{p}\sum_{j=1}^{p} y_{.j..}^2 - \frac{y_{....}^2}{N}$	$p-1$
Tratamientos con letras griegas	$SS_G = \frac{1}{p}\sum_{k=1}^{p} y_{..k.}^2 - \frac{y_{....}^2}{N}$	$p-1$
Renglones	$SS_{\text{Renglones}} = \frac{1}{p}\sum_{i=1}^{p} y_{i...}^2 - \frac{y_{....}^2}{N}$	$p-1$
Columnas	$SS_{\text{Columnas}} = \frac{1}{p}\sum_{l=1}^{p} y_{...l}^2 - \frac{y_{....}^2}{N}$	$p-1$
Error	SS_E (por sustracción)	$(p-3)(p-1)$
Total	$SS_T = \sum_i\sum_j\sum_k\sum_l y_{ijkl}^2 - \frac{y_{....}^2}{N}$	p^2-1

El modelo estadístico para el diseño de cuadrado grecolatino es

$$y_{ijkl} = \mu + \theta_i + \tau_j + \omega_k + \Psi_l + \varepsilon_{ijkl} \begin{cases} i = 1, 2, \ldots, p \\ j = 1, 2, \ldots, p \\ k = 1, 2, \ldots, p \\ l = 1, 2, \ldots, p \end{cases} \tag{4-25}$$

donde y_{ijkl} es la observación del renglón i y la columna l para la letra latina j y la letra griega k, θ_i es el efecto del renglón i-ésimo, τ_j es el efecto del tratamiento de letra latina j, ω_k es el efecto del tratamiento de letra griega k, Ψ_l es el efecto de la columna l, y ε_{ijkl} es un componente NID$(0, \sigma^2)$ del error aleatorio. Sólo son necesarios dos de los cuatro subíndices para identificar completamente una observación.

El análisis de varianza es muy parecido al de un cuadrado latino. Puesto que las letras griegas aparecen exactamente una vez en cada renglón y columna, y exactamente una vez con cada letra latina, el factor representado por las letras griegas es ortogonal a los renglones, las columnas y los tratamientos de letras latinas. Por lo tanto, puede calcularse una suma de cuadrados debida al factor de las letras griegas a partir de los totales de las letras griegas y el error experimental se reduce adicionalmente en esta cantidad. En la tabla 4-19 se ilustran los detalles de los cálculos. La hipótesis nula de la igualdad de tratamientos de renglones, columnas, letras latinas y letras griegas, se probaría dividiendo el cuadrado medio correspondiente por el cuadrado medio del error. La región de rechazo es la cola superior del punto de la distribución $F_{p-1, (p-3)(p-1)}$.

EJEMPLO 4-4 ··

Suponga que en el experimento de la carga propulsora del ejemplo 4-3 un factor adicional, los montajes de prueba, podría ser importante. Sea que haya cinco montajes de prueba denotados por las letras griegas α, β, γ, δ y ε. En la tabla 4-20 se muestra el diseño de cuadrado grecolatino 5 × 5 resultante.

Observe que, debido a que los totales de los lotes de materia prima (renglones), los operadores (columnas) y las formulaciones (letras latinas) son idénticos a los del ejemplo 4-3, se tiene

$$SS_{\text{Lotes}} = 68.00 \qquad SS_{\text{Operadores}} = 150.00 \qquad \text{y} \qquad SS_{\text{Formulaciones}} = 330.00$$

Tabla 4-20 Diseño del cuadrado grecolatino para el problema de la carga propulsora

Lotes de materia prima	Operadores 1	2	3	4	5	$y_{i...}$
1	$A\alpha = -1$	$B\gamma = -5$	$C\varepsilon = -6$	$D\beta = -1$	$E\delta = -1$	-14
2	$B\beta = -8$	$C\delta = -1$	$D\alpha = 5$	$E\gamma = 2$	$A\varepsilon = 11$	9
3	$C\gamma = -7$	$D\varepsilon = 13$	$E\beta = 1$	$A\delta = 2$	$B\alpha = -4$	5
4	$D\delta = 1$	$E\alpha = 6$	$A\gamma = 1$	$B\varepsilon = -2$	$C\beta = -3$	3
5	$E\varepsilon = -3$	$A\beta = 5$	$B\delta = -5$	$C\alpha = 4$	$D\gamma = 6$	7
$y_{..l}$	-18	18	-4	5	9	$10 = y_{....}$

Los totales de los montajes de prueba (las letras griegas) son

Letra griega	**Total de la prueba de ensamblaje**
α	$y_{..1.} = 10$
β	$y_{..2.} = -6$
γ	$y_{..3.} = -3$
δ	$y_{..4.} = -4$
ε	$y_{..5.} = 13$

Por lo tanto, la suma de cuadrados debida a los montajes de prueba es

$$SS_{\text{Ensamblajes}} = \frac{1}{p}\sum_{k=1}^{p} y_{..k.}^2 - \frac{y_{....}^2}{N}$$

$$= \frac{1}{5}[10^2 + (-6)^2 + (-3)^2 + (-4)^2 + 13^2] - \frac{(10)^2}{25} = 62.00$$

En la tabla 4-21 se resume el análisis de varianza completo. Las formulaciones son diferentes significativamente en 1%. Al comparar las tablas 4-21 y 4-12, se observa que al sacar la variabilidad debida a los montajes de prueba, el error experimental disminuye. Sin embargo, al disminuir el error experimental, se han reducido también los grados de libertad de 12 (en el diseño del cuadrado latino del ejemplo 4-3) a 8. Por lo tanto, la estimación del error tiene menos grados de libertad, y la prueba puede ser menos sensible.

Tabla 4-21 Análisis de varianza del problema de la carga propulsora

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0	Valor P
Formulaciones	330.00	4	82.50	10.00	0.0033
Lotes de materia prima	68.00	4	17.00		
Operadores	150.00	4	37.50		
Montajes de la prueba	62.00	4	15.50		
Error	66.00	8	8.25		
Total	676.00	24			

Puede hacerse cierta ampliación del concepto de los pares ortogonales de cuadrados latinos que forman un cuadrado grecolatino. Un **hipercuadrado** $p \times p$ es un diseño en el que se superponen tres o más cuadrados latinos ortogonales $p \times p$. En general, hasta $p + 1$ factores podrían estudiarse si se dispone de un conjunto completo de $p - 1$ cuadrados latinos ortogonales. En este diseño se utilizarían todos los $(p + 1)(p - 1) = p^2 - 1$ grados de libertad, por lo que se necesita una estimación independiente de la varianza del error. Desde luego, no debe haber interacciones entre los factores cuando se usan hipercuadrados.

4-4 DISEÑOS DE BLOQUES INCOMPLETOS BALANCEADOS

En ciertos experimentos en los que se utilizan diseños de bloques aleatorizados quizá no sea posible correr todas las combinaciones de los tratamientos en cada bloque. Situaciones como ésta ocurren generalmente por limitaciones del aparato experimental o de las instalaciones o por el tamaño físico del bloque. Por ejemplo, en el experimento de la prueba de la dureza (ejemplo 4-1), suponga que debido a sus dimensiones cada ejemplar de prueba sólo puede usarse para probar tres puntas. Por lo tanto, no es posible probar todas las puntas en cada uno de los ejemplares. Para este tipo de problema es posible utilizar diseños de bloques aleatorizados en los que cada tratamiento no está presente en cada bloque. Estos diseños se conocen como **diseños de bloques incompletos aleatorizados.**

Cuando las comparaciones de todos los tratamientos son igualmente importantes, las combinaciones de los tratamientos usadas en cada bloque deberán seleccionarse en una forma balanceada, es decir, de tal manera que cualquier par de tratamientos ocurra conjuntamente el mismo número de veces que cualquier otro par. Por lo tanto, un **diseño de bloques incompletos balanceados (BIBD, *balanced incomplete block design*)** es un diseño de bloques incompletos en el que dos tratamientos cualesquiera aparecen conjuntamente el mismo número de veces. Suponga que hay a tratamientos y que cada bloque puede contener exactamente k $(k < a)$ tratamientos. Un diseño de bloques incompletos balanceados puede construirse tomando $\binom{a}{k}$ bloques y asignando una combinación de tratamientos diferente a cada bloque. Con frecuencia, sin embargo, puede obtenerse un diseño balanceado con menos de $\binom{a}{k}$ bloques. Tablas de BIBD se proporcionan en Fisher y Yates [45], Davies [36] y Cochran y Cox [26].

Como un ejemplo, suponga que un ingeniero químico piensa que el tiempo de reacción de un proceso químico es una función del tipo de catalizador empleado. Se están investigando cuatro catalizadores. El procedimiento experimental consiste en seleccionar un lote de materia prima, cargar la planta piloto, aplicar cada catalizador en una corrida separada de la planta piloto y observar el tiempo de reacción. Debido a que las variaciones en los lotes de materia prima pueden afectar el desempeño de los catalizadores, el ingeniero decide usar los lotes de materia prima como bloques. Sin embargo, cada lote es apenas lo suficientemente grande para permitir que se prueben tres catalizadores. Por lo tanto, debe usarse un diseño

Tabla 4-22 Diseño de bloques incompletos balanceados para el experimento del catalizador

Tratamiento	Bloque (lote de materia prima)				
(catalizador)	1	2	3	4	$y_{i.}$
1	73	74	–	71	218
2	–	75	67	72	214
3	73	75	68	–	216
4	75	–	72	75	222
$y_{.j}$	221	224	207	218	$870 = y_{..}$

de bloques incompletos aleatorizados. El diseño de bloques incompletos balanceados para este experimento, junto con las observaciones registradas, se muestran en la tabla 4-22. El orden en que se corren los catalizadores en cada bloque está aleatorizado.

4-4.1 Análisis estadístico del diseño de bloques incompletos balanceados

Como de costumbre, se supone que hay a tratamientos y b bloques. Además, se supone que cada bloque contiene k tratamientos, que cada tratamiento ocurre r veces en el diseño (o que se hacen r réplicas del mismo), y que hay $N = ar = bk$ observaciones en total. Asimismo, el número de veces que cada par de tratamientos aparece en el mismo bloque es

$$\lambda = \frac{r(k-1)}{a-1}$$

Si $a = b$, se dice que el diseño es **simétrico.**

El parámetro λ debe ser un entero. Para deducir la relación para λ, considere cualquier tratamiento, por ejemplo el tratamiento 1. Puesto que el tratamiento 1 aparece en r bloques y hay otros $k-1$ tratamientos en cada uno de esos bloques, hay $r(k-1)$ observaciones en un bloque que contiene al tratamiento 1. Estas $r(k-1)$ observaciones también tienen que representar a los $a-1$ tratamientos restantes λ veces. Por lo tanto, $\lambda(a-1) = r(k-1)$.

El **modelo estadístico** del BIBD es

$$y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij} \tag{4-26}$$

donde y_{ij} es la observación i-ésima en el bloque j-ésimo, μ es la media global, τ_i es el efecto del tratamiento i-ésimo, β_j es el efecto del bloque j-ésimo, y ε_{ij} es el componente NID$(0, \sigma^2)$ del error aleatorio. La variabilidad total en los datos se expresa por la suma de cuadrados totales corregida:

$$SS_T = \sum_i \sum_j y_{ij}^2 - \frac{y_{..}^2}{N} \tag{4-27}$$

Puede hacerse la partición de la variabilidad total en

$$SS_T = SS_{\text{Tratamientos(ajustados)}} + SS_{\text{Bloques}} + SS_E$$

donde la suma de cuadrados de los tratamientos está **ajustada** para separar los efectos de los tratamientos y de los bloques. Este ajuste es necesario porque cada tratamiento está representado en un conjunto diferente de r bloques. Por lo tanto, las diferencias entre los totales de los tratamientos no ajustados $y_{1.}, y_{2.}, \ldots, y_{a.}$ también son afectadas por las diferencias entre los bloques.

La suma de cuadrados de los bloques es

$$SS_{\text{Bloques}} = \frac{1}{k}\sum_{j=1}^{b} y_{.j}^2 - \frac{y_{..}^2}{N} \tag{4-28}$$

donde $y_{.j}$ es el total del bloque j-ésimo. SS_{Bloques} tiene $b-1$ grados de libertad. La suma de cuadrados de los tratamientos ajustada es

$$SS_{\text{Tratamientos(ajustados)}} = \frac{k\sum_{i=1}^{a} Q_i^2}{\lambda a} \tag{4-29}$$

Tabla 4-23 Análisis de varianza del diseño de bloques incompletos balanceados

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0
Tratamientos (ajustados)	$\frac{k\sum Q_i^2}{\lambda a}$	$a-1$	$\frac{SS_{\text{Tratamientos(ajustados)}}}{a-1}$	$F_0=\frac{MS_{\text{Tratamientos(ajustados)}}}{MS_E}$
Bloques	$\frac{1}{k}\sum y_{.j}^2-\frac{y_{..}^2}{N}$	$b-1$	$\frac{SS_{\text{Bloques}}}{b-1}$	
Error	SS_E (por sustracción)	$N-a-b+1$	$\frac{SS_E}{N-a-b+1}$	
Total	$\sum\sum y_{ij}^2-\frac{y_{..}^2}{N}$	$N-1$		

donde Q_i es el total ajustado del tratamiento i-ésimo, el cual se calcula como

$$Q_i = y_{i.} - \frac{1}{k}\sum_{j=1}^{b} n_{ij}y_{.j} \qquad i=1,2,\ldots,a \tag{4-30}$$

con $n_{ij}=1$ si el tratamiento i aparece en el bloque j y $n_{ij}=0$ en caso contrario. Los totales de los tratamientos ajustados siempre sumarán cero. $SS_{\text{Tratamientos(ajustados)}}$ tiene $a-1$ grados de libertad. La suma de cuadrados del error se calcula por sustracción como

$$SS_E = SS_T - SS_{\text{Tratamientos(ajustados)}} - SS_{\text{Bloques}} \tag{4-31}$$

y tiene $N-a-b+1$ grados de libertad.

El estadístico apropiado para probar la igualdad de los efectos de los tratamientos es

$$F_0 = \frac{MS_{\text{Tratamientos(ajustados)}}}{MS_E}$$

En la tabla 4-23 se resume el análisis de varianza.

EJEMPLO 4-5

Considere los datos de la tabla 4-22 para el experimento del catalizador. Se trata de un BIBD con $a=4$, $b=4$, $k=3$, $r=3$, $\lambda=2$ y $N=12$. El análisis de estos datos es el siguiente. La suma de cuadrados total es

$$SS_T = \sum_i \sum_j y_{ij}^2 - \frac{y_{..}^2}{12}$$

$$= 63{,}156 - \frac{(870)^2}{12} = 81.00$$

La suma de cuadrados de los bloques se encuentra con la ecuación 4-28 como

$$SS_{\text{Bloques}} = \frac{1}{3}\sum_{j=1}^{4} y_{.j}^2 - \frac{y_{..}^2}{12}$$

$$= \frac{1}{3}[(221)^2+(207)^2+(224)^2+(218)^2] - \frac{(870)^2}{12} = 55.00$$

Tabla 4-24 Análisis de varianza del ejemplo 4-5

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0	Valor P
Tratamientos (ajustados para los bloques)	22.75	3	7.58	11.66	0.0107
Bloques	55.00	3	–		
Error	3.25	5	0.65		
Total	81.00	11			

Para calcular la suma de cuadrados de los tratamientos ajustados para los bloques, primero se determinan los totales de los tratamientos ajustados utilizando la ecuación 4-30 como

$$Q_1 = (218) - \tfrac{1}{3}(221 + 224 + 218) = -9/3$$
$$Q_2 = (214) - \tfrac{1}{3}(207 + 224 + 218) = -7/3$$
$$Q_3 = (216) - \tfrac{1}{3}(221 + 207 + 224) = -4/3$$
$$Q_4 = (222) - \tfrac{1}{3}(221 + 207 + 218) = 20/3$$

La suma de cuadrados de los tratamientos ajustados se calcula con la ecuación 4-29 como

$$SS_{\text{Tratamientos(ajustados)}} = \frac{k\sum_{i=1}^{4} Q_i^2}{\lambda a}$$
$$= \frac{3[(-9/3)^2 + (-7/3)^2 + (-4/3)^2 + (20/3)^2]}{(2)(4)} = 22.75$$

La suma de cuadrados del error se obtiene por sustracción como

$$SS_E = SS_T - SS_{\text{Tratamientos(ajustados)}} - SS_{\text{Bloques}}$$
$$= 81.00 - 22.75 - 55.00 = 3.25$$

En la tabla 4-24 se muestra el análisis de varianza. Puesto que el valor P es pequeño, se concluye que el catalizador empleado tiene un efecto significativo sobre el tiempo de reacción.

Si el factor bajo estudio es fijo, las pruebas para las medias de tratamientos individuales pueden ser de interés. Si se emplean contrastes ortogonales, los contrastes deben hacerse sobre los **totales de los tratamientos ajustados,** las $\{Q_i\}$ en lugar de las $\{y_{i.}\}$. La suma de cuadrados de los contrastes es

$$SS_C = \frac{k\left(\sum_{i=1}^{a} c_i Q_i\right)^2}{\lambda a \sum_{i=1}^{a} c_i^2}$$

donde $\{c_i\}$ son los coeficientes de los contrastes. Pueden usarse otros métodos de comparación múltiple

para comparar todos los pares de efectos de los tratamientos ajustados (sección 4-4.2), los cuales se estiman con $\hat{\tau}_i = kQ_i/(\lambda a)$. El error estándar del efecto de un tratamiento ajustado es

$$S = \sqrt{\frac{kMS_E}{\lambda a}} \tag{4-32}$$

En el análisis que acaba de describirse, se ha hecho la partición de la suma de cuadrados total en una suma de cuadrados de los tratamientos ajustados, una suma de cuadrados de los bloques sin ajuste y una suma de cuadrados del error. En ocasiones habría interés en evaluar los efectos de los bloques. Para ello se requiere hacer una partición alternativa de SS_T, es decir,

$$SS_T = SS_{\text{Tratamientos}} + SS_{\text{Bloques(ajustados)}} + SS_E$$

Aquí, $SS_{\text{Tratamientos}}$ está sin ajuste. Si el diseño es simétrico, es decir, si $a = b$, puede obtenerse una fórmula simple para $SS_{\text{Bloques(ajustados)}}$. Los totales de los bloques ajustados son

$$Q'_j = y_{.j} - \frac{1}{r}\sum_{i=1}^{a} n_{ij} y_{i.} \qquad j = 1, 2, \ldots, b \tag{4-33}$$

y

$$SS_{\text{Bloques(ajustados)}} = \frac{r\sum_{j=1}^{b} (Q'_j)^2}{\lambda b} \tag{4-34}$$

El BIBD del ejemplo 4-5 es simétrico porque $a = b = 4$. Por lo tanto,

$$Q'_1 = (221) - \tfrac{1}{3}(218+216+222) = 7/3$$
$$Q'_2 = (224) - \tfrac{1}{3}(218+214+216) = 24/3$$
$$Q'_3 = (207) - \tfrac{1}{3}(214+216+222) = -31/3$$
$$Q'_4 = (218) - \tfrac{1}{3}(218+214+222) = 0$$

y

$$SS_{\text{Bloques(ajustados)}} = \frac{3[(7/3)^2 + (24/3)^2 + (-31/3)^2 + (0)^2]}{(2)(4)} = 66.08$$

Asimismo,

$$SS_{\text{Tratamientos}} = \frac{(218)^2 + (214)^2 + (216)^2 + (222)^2}{3} - \frac{(870)^2}{12} = 11.67$$

Tabla 4-25 Análisis de varianza del ejemplo 4-5, incluyendo tanto los tratamientos como los bloques

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0	Valor P
Tratamientos (ajustados)	22.75	3	7.58	11.66	0.0107
Tratamientos (sin ajuste)	11.67	3			
Bloques (sin ajuste)	55.00	3			
Bloques (ajustados)	66.08	3	22.03	33.90	0.0010
Error	3.25	5	0.65		
Total	81.0	11			

En la tabla 4-25 se presenta un resumen del análisis de varianza del BIBD simétrico. Observe que las sumas de cuadrados asociadas con los cuadrados medios de la tabla 4-25 no producen la suma de cuadrados total, es decir,

$$SS_T \neq SS_{\text{Tratamientos(ajustados)}} + SS_{\text{Bloques(ajustados)}} + SS_E$$

Esto es consecuencia del carácter no ortogonal de los tratamientos y los bloques.

Salida de computadora

Existen varios paquetes de computadora que realizarán el análisis de un diseño de bloques incompletos balanceados. El procedimiento de Modelos Lineales Generales (General Linear Models) del SAS es uno de ellos, y Minitab, un paquete de estadística para computadoras personales de uso generalizado, es otro. La parte superior de la tabla 4-26 es la salida del procedimiento de Modelos Lineales Generales de Minitab para el ejemplo 4-5. Al comparar las tablas 4-26 y 4-25, se observa que Minitab ha calculado la suma de cuadrados de los tratamientos ajustados y la suma de cuadrados de los bloques ajustados (en la salida de Minitab se les llama "AdjSS" o SS ajustada).

La parte inferior de la tabla 4-26 es un análisis de comparaciones múltiples, en el que se utiliza el método de Tukey. Se presentan los intervalos de confianza para las diferencias de todos los pares de medias y la prueba de Tukey. Observe que el método de Tukey llevaría a la conclusión de que el catalizador 4 es diferente de los otros tres.

4-4.2 Estimación de mínimos cuadrados de los parámetros

Considere la estimación de los efectos de los tratamientos en el modelo BIBD. Las ecuaciones normales de mínimos cuadrados son

$$\mu: N\hat{\mu} + r\sum_{i=1}^{a} \hat{\tau}_i + k\sum_{j=1}^{b} \hat{\beta}_j = y_{..}$$

$$\tau_i: r\hat{\mu} + r\hat{\tau}_i + \sum_{j=1}^{b} n_{ij}\hat{\beta}_j = y_{i.} \qquad i = 1, 2, \ldots, a \tag{4-35}$$

$$\beta_j: k\hat{\mu} + \sum_{i=1}^{a} n_{ij}\hat{\tau}_i + k\hat{\beta}_j = y_{.j} \qquad j = 1, 2, \ldots, b$$

Al imponer las restricciones $\Sigma\hat{\tau}_i = \Sigma\hat{\beta}_j = 0$, se encuentra que $\hat{\mu} = \bar{y}_{..}$. Además, al utilizar las ecuaciones para $\{\beta_j\}$ para eliminar los efectos de los bloques de las ecuaciones para $\{\tau_i\}$, se obtiene

$$rk\hat{\tau}_i - r\hat{\tau}_i - \sum_{j=1}^{b}\sum_{\substack{p=1\\p\neq i}}^{a} n_{ij}n_{pj}\hat{\tau}_p = ky_{i.} - \sum_{j=1}^{b} n_{ij}y_{.j} \tag{4-36}$$

Observe que el miembro del lado derecho de la ecuación 4-36 es kQ_i, donde Q_i es el total del tratamiento ajustado i-ésimo (ver la ecuación 4-29). Entonces, puesto que $\Sigma_{j=1}^{b} n_{ij}n_{pj} = \lambda$ si $p \neq i$ y $n_{pj}^2 = n_{pj}$ (ya que $n_{pj} =$ 0 o 1), la ecuación 4-36 puede reescribirse como

$$r(k-1)\hat{\tau}_i - \lambda\sum_{\substack{p=1\\p\neq i}}^{a} \hat{\tau}_p = kQ_i \qquad i = 1, 2, \ldots, a \tag{4-37}$$

Tabla 4-26 Análisis de Minitab (Modelo Lineal General) para el ejemplo 4-5

Modelo Lineal General

```
Factor      Type  Levels Values
Catalyst   fixed       4 1 2 3 4
Block      fixed       4 1 2 3 4

Analysis of Variance for Time, using Adjusted SS for Tests

Source     DF     Seq SS     Adj SS     Adj MS       F      P
Catalyst    3     11.667     22.750      7.583   11.67  0.011
Block       3     66.083     66.083     22.028   33.89  0.001
Error       5      3.250      3.250      0.650
Total      11     81.000

Tukey 95.0% Simultaneous Confidence Intervals
Response Variable Time
All Pairwise Comparisons among Levels of Catalyst

Catalyst = 1 subtracted from:

Catalyst     Lower    Center    Upper   ----------+---------+---------+------
2           -2.327    0.2500    2.827   (---------*---------)
3           -1.952    0.6250    3.202    (---------*---------)
4            1.048    3.6250    6.202               (---------*---------)
                                         ----------+---------+---------+------
                                                 0.0       2.5       5.0


Catalyst = 2 subtracted from:

Catalyst     Lower    Center     Upper  ----------+---------+---------+------
3           -2.202    0.3750     2.952   (---------*---------)
4            0.798    3.3750     5.952             (---------*---------)
                                        ----------+---------+---------+------
                                                 0.0       2.5       5.0


Catalyst = 3 subtracted from:

Catalyst     Lower    Center     Upper  ----------+---------+---------+------
4           0.4228     3.000     5.577            (---------*---------)
                                        ----------+---------+---------+------
                                                 0.0       2.5       5.0


Tukey Simultaneous Tests
Response Variable Time
All Pairwise Comparisons among Levels of Catalyst

Catalyst = 1 subtracted from:

Level        Difference       SE of              Adjusted
Catalyst       of Means  Difference   T-Value     P-Value
2                0.2500      0.6982    0.3581      0.9825
3                0.6250      0.6982    0.8951      0.8085
4                3.6250      0.6982    5.1918      0.0130

Catalyst = 2 subtracted from:

Level        Difference       SE of              Adjusted
Catalyst       of Means  Difference   T-Value     P-Value
3                0.3750      0.6982    0.5371      0.9462
4                3.3750      0.6982    4.8338      0.0175

Catalyst = 3 subtracted from:

Level        Difference       SE of              Adjusted
Catalyst       of Means  Difference   T-Value     P-Value
4                 3.000      0.6982     4.297      0.0281
```

Por último, observe que la restricción $\Sigma_{i=1}^{a}\hat{\tau}_i = 0$ implica que $\Sigma_{\substack{p=1\\p\neq i}}^{a}\hat{\tau}_p = -\hat{\tau}_i$ y recuerde que $r(k-1) = \lambda(a-1)$, de donde se obtiene

$$\lambda a\hat{\tau}_i = kQ_i \qquad i = 1, 2, \ldots, a \tag{4-38}$$

Por lo tanto, los estimadores de mínimos cuadrados de los efectos de los tratamientos en el modelo de bloques incompletos balanceados son

$$\hat{\tau}_i = \frac{kQ_i}{\lambda a} \qquad i = 1, 2, \ldots, a \tag{4-39}$$

Como una ilustración, considere el BIBD del ejemplo 4-5. Puesto que $Q_1 = -9/3$, $Q_2 = -7/3$, $Q_3 = -4/3$ y $Q_4 = 20/3$, se obtiene

$$\hat{\tau}_1 = \frac{3(-9/3)}{(2)(4)} = -9/8 \qquad \hat{\tau}_2 = \frac{3(-7/3)}{(2)(4)} = -7/8$$
$$\hat{\tau}_3 = \frac{3(-4/3)}{(2)(4)} = -4/8 \qquad \hat{\tau}_4 = \frac{3(20/3)}{(2)(4)} = 20/8$$

como se encontró en la sección 4-4.1.

4-4.3 Recuperación de información interbloques en el diseño de bloques incompletos balanceados

Al análisis del BIBD presentado en la sección 4-4.1 suele llamársele el **análisis intrabloques** porque las diferencias de los bloques se eliminan y todos los contrastes de los efectos de los tratamientos pueden expresarse como comparaciones entre las observaciones del mismo bloque. Este análisis es apropiado independientemente de si los bloques son fijos o aleatorios. Yates [113c] señaló que si los efectos de los bloques son variables aleatorias no correlacionadas con medias cero y varianza σ_β^2, es posible obtener información adicional acerca de los efectos de los tratamientos τ_i. Yates llamó **análisis interbloques** al método para obtener esta información adicional.

Considere los totales de los bloques $y_{.j}$ como una colección de b observaciones. El modelo para estas observaciones (siguiendo a John [61d]) es

$$y_{.j} = k\mu + \sum_{i=1}^{a} n_{ij}\tau_i + \left(k\beta_j + \sum_{i=1}^{a} \varepsilon_{ij}\right) \tag{4-40}$$

donde el término entre paréntesis puede considerarse como el error. Los estimadores interbloques de μ y τ_i se encuentran minimizando la función de mínimos cuadrados

$$L = \sum_{j=1}^{b}\left(y_{.j} - k\mu - \sum_{i=1}^{a} n_{ij}\tau_i\right)^2$$

Se obtienen así las siguientes ecuaciones normales de mínimos cuadrados:

$$\begin{aligned}
&\mu: N\tilde{\mu} + r\sum_{i=1}^{a}\tilde{\tau}_i = y_{..}\\
&\tau_i: kr\tilde{\mu} + r\tilde{\tau}_i + \lambda\sum_{\substack{p=1\\p\neq i}}^{a}\tilde{\tau}_p = \sum_{j=1}^{b} n_{ij}y_{.j} \qquad i = 1, 2, \ldots, a
\end{aligned} \tag{4-41}$$

donde $\tilde{\mu}$ y $\tilde{\tau}_i$ denotan los **estimadores interbloques.** Al imponer la restricción $\Sigma_{i=1}^{a}\hat{\tau}_i = 0$, se obtienen las soluciones de las ecuaciones 4-41 como

$$\tilde{\mu} = \bar{y}_{..} \tag{4-42}$$

$$\tilde{\tau}_i = \frac{\sum_{j=1}^{b} n_{ij} y_{.j} - kr\bar{y}_{..}}{r-\lambda} \qquad i = 1, 2, \ldots, a \tag{4-43}$$

Es posible demostrar que los estimadores interbloques $\{\tilde{\tau}_i\}$ y los estimadores intrabloques $\{\hat{\tau}_i\}$ no están correlacionados.

Los estimadores interbloques $\{\tilde{\tau}_i\}$ pueden diferir de los estimadores intrabloques $\{\hat{\tau}_i\}$. Por ejemplo, los estimadores interbloques para el BIBD del ejemplo 4-5 se calculan de la siguiente manera:

$$\tilde{\tau}_1 = \frac{663-(3)(3)(72.50)}{3-2} = 10.50$$

$$\tilde{\tau}_2 = \frac{649-(3)(3)(72.50)}{3-2} = -3.50$$

$$\tilde{\tau}_3 = \frac{652-(3)(3)(72.50)}{3-2} = -0.50$$

$$\tilde{\tau}_4 = \frac{646-(3)(3)(72.50)}{3-2} = -6.50$$

Observe que los valores de $\Sigma_{j=1}^{b} n_{ij} y_{.j}$ se usaron en la página 157 para calcular los totales de los tratamientos ajustados en el análisis intrabloques.

Suponga ahora que quieren combinarse los estimadores interbloques e intrabloques para obtener una sola estimación de la varianza mínima insesgada de cada τ_i. Es posible demostrar que $\hat{\tau}_i$ y $\tilde{\tau}_i$ son insesgados y también que

$$V(\hat{\tau}_i) = \frac{k(a-1)}{\lambda a^2}\sigma^2 \qquad \text{(intrabloques)}$$

y

$$V(\tilde{\tau}_i) = \frac{k(a-1)}{a(r-\lambda)}(\sigma^2 + k\sigma_\beta^2) \qquad \text{(intrabloques)}$$

Se usa una combinación lineal de los dos estimadores, por ejemplo

$$\tau_i^* = \alpha_1\hat{\tau}_i + \alpha_2\hat{\tau}_i \tag{4-44}$$

para estimar τ_i. En este método de estimación, el estimador combinado insesgado de la varianza mínima τ_i^* deberá tener las ponderaciones $\alpha_1 = u_1/(u_1 + u_2)$ y $\alpha_2 = u_2/(u_1 + u_2)$, donde $u_1 = 1/V(\hat{\tau}_i)$ y $u_2 = 1/V(\tilde{\tau}_i)$. Por lo tanto, las ponderaciones óptimas son inversamente proporcionales a las varianzas de $\hat{\tau}_i$ y $\tilde{\tau}_i$. Esto implica que el mejor estimador combinado es

$$\tau_i^* = \frac{\hat{\tau}_i \dfrac{k(a-1)}{a(r-\lambda)}(\sigma^2 + k\sigma_\beta^2) + \tilde{\tau}_i \dfrac{k(a-1)}{\lambda a^2}\sigma^2}{\dfrac{k(a-1)}{\lambda a^2}\sigma^2 + \dfrac{k(a-1)}{a(r-\lambda)}(\sigma^2 + k\sigma_\beta^2)} \qquad i = 1, 2, \ldots, a$$

que puede simplificarse como

$$\tau_i^* = \frac{kQ_i(\sigma^2 + k\sigma_\beta^2) + \left(\sum_{j=1}^{b} n_{ij} y_{.j} - kr\bar{y}_{..}\right)\sigma^2}{(r-\lambda)\sigma^2 + \lambda a(\sigma^2 + k\sigma_\beta^2)} \qquad i = 1, 2, \ldots, a \tag{4-45}$$

Desafortunadamente, la ecuación 4-45 no puede usarse para estimar τ_i porque no se conocen las varianzas σ^2 y σ_β^2. El enfoque común es estimar σ^2 y σ_β^2 a partir de los datos y sustituir estos parámetros de la ecuación 4-45 con las estimaciones. La estimación que suele tomarse para σ^2 es el cuadrado medio del error del análisis de varianza intrabloques, o el **error intrabloques.** Por lo tanto,

$$\hat{\sigma}^2 = MS_E$$

La estimación de σ_β^2 se encuentra a partir del cuadrado medio de los bloques ajustados para los tratamientos. En general, para un diseño de bloques incompletos balanceados, este cuadrado medio es

$$MS_{\text{Bloques(ajustados)}} = \frac{\left(\dfrac{k\sum_{i=1}^{a} Q_i^2}{\lambda a} + \sum_{j=1}^{b} \dfrac{y_{.j}^2}{k} - \sum_{i=1}^{a} \dfrac{y_{i.}^2}{r}\right)}{(b-1)} \tag{4-46}$$

y su valor esperado (cuya deducción se hace en Graybill [50]) es

$$E[MS_{\text{Bloques(ajustados)}}] = \sigma^2 + \frac{a(r-1)}{b-1}\sigma_\beta^2$$

Por lo tanto, si $MS_{\text{Bloques(ajustados)}} > MS_E$, la estimación de $\hat{\sigma}_\beta^2$ es

$$\hat{\sigma}_\beta^2 = \frac{[MS_{\text{Bloques(ajustados)}} - MS_E](b-1)}{a(r-1)} \tag{4-47}$$

y si $MS_{\text{Bloques(ajustados)}} \leq MS_E$, se hace $\hat{\sigma}_\beta^2 = 0$. Esto resulta en el estimador combinado

$$\tau_i^* = \begin{cases} \dfrac{kQ_i(\hat{\sigma}^2 + k\hat{\sigma}_\beta^2) + \left(\sum_{j=1}^{b} n_{ij} y_{.j} - kr\bar{y}_{..}\right)\hat{\sigma}^2}{(r-\lambda)\hat{\sigma}^2 + \lambda a(\hat{\sigma}^2 + k\hat{\sigma}_\beta^2)}, & \hat{\sigma}_\beta^2 > 0 \quad (4-48\text{a}) \\[2ex] \dfrac{y_{i.} - (1/a)y_{..}}{r}, & \hat{\sigma}_\beta^2 = 0 \quad (4-48\text{b}) \end{cases}$$

A continuación se calculan las estimaciones combinadas para los datos del ejemplo 4-5. Por la tabla 4-25 se obtiene $\hat{\sigma}^2 = MS_E = 0.65$ y $MS_{\text{Bloques(ajustados)}} = 22.03$. (Observe que para calcular $MS_{\text{Bloques(ajustados)}}$ se hace uso del hecho de que éste es un diseño simétrico. En general, debe usarse la ecuación 4-46.) Puesto que $MS_{\text{Bloques(ajustados)}} > MS_E$, se usa la ecuación 4-47 para estimar σ_β^2 como

$$\hat{\sigma}_\beta^2 = \frac{(22.03 - 0.65)(3)}{4(3-1)} = 8.02$$

Por lo tanto, pueden sustituirse $\hat{\sigma}^2 = 0.65$ y $\hat{\sigma}_\beta^2 = 8.02$ en la ecuación 4-48a para obtener las estimaciones combinadas que se enlistan enseguida. Por conveniencia, también se presentan las estimaciones intrablo-

ques e interbloques. En este ejemplo, las estimaciones combinadas están próximas a las estimaciones intrabloques debido a que la varianza de las estimaciones interbloques es relativamente grande.

Parámetro	Estimación intrabloques	Estimación interbloques	Estimación combinada
τ_1	−1.12	10.50	−1.09
τ_2	−0.88	−3.50	−0.88
τ_3	−0.50	−0.50	−0.50
τ_4	2.50	−6.50	2.47

4-5 PROBLEMAS

4-1. Un químico quiere probar el efecto de cuatro agentes químicos sobre la resistencia de un tipo particular de tela. Debido a que podría haber variabilidad de un rollo de tela a otro, el químico decide usar un diseño de bloques aleatorizados, con los rollos de tela considerados como bloques. Selecciona cinco rollos y aplica los cuatro agentes químicos de manera aleatoria a cada rollo. A continuación se presentan las resistencias a la tensión resultantes. Analizar los datos de este experimento (utilizar $\alpha = 0.05$) y sacar las conclusiones apropiadas.

Agente químico	Rollo 1	2	3	4	5
1	73	68	74	71	67
2	73	67	75	72	70
3	75	68	78	73	68
4	73	71	75	75	69

4-2. Se están comparando tres soluciones de lavado diferentes a fin de estudiar su efectividad para retardar el crecimiento de bacterias en contenedores de leche de 5 galones. El análisis se hace en un laboratorio y sólo pueden realizarse tres ensayos en un día. Puesto que los días podrían representar una fuente potencial de variabilidad, el experimentador decide usar un diseño de bloques aleatorizados. Se hacen observaciones en cuatro días, cuyos datos se muestran enseguida. Analizar los datos de este experimento (utilizar $\alpha = 0.05$) y sacar las conclusiones apropiadas.

Solución	Días 1	2	3	4
1	13	22	18	39
2	16	24	17	44
3	5	4	1	22

4-3. Graficar las resistencias a la tensión medias observadas para cada tipo de agente químico en el problema 4-1 y compararlas con una distribución t con la escalación apropiada. ¿Qué conclusiones se sacarían a partir de esta representación gráfica?

4-4. Graficar los conteos de bacterias promedio para cada solución en el problema 4-2 y compararlos con una distribución t escalada. ¿Qué conclusiones pueden sacarse?

4-5. En un artículo de *Fire Safety Journal* ("El efecto del diseño de boquillas en la estabilidad y el desempeño de surtidores de agua turbulenta", vol. 4) se describe un experimento en el que se determinó un factor de la forma para varios diseños diferentes de boquillas con seis niveles de la velocidad del flujo de salida del surtidor.

El interés se centró en las diferencias potenciales entre los diseños de las boquillas, con la velocidad considerada como una variable perturbadora. Los datos se presentan a continuación.

Diseño de la boquilla	Velocidad del flujo de salida del surtidor (m/s)					
	11.73	14.37	16.59	20.43	23.46	28.74
1	0.78	0.80	0.81	0.75	0.77	0.78
2	0.85	0.85	0.92	0.86	0.81	0.83
3	0.93	0.92	0.95	0.89	0.89	0.83
4	1.14	0.97	0.98	0.88	0.86	0.83
5	0.97	0.86	0.78	0.76	0.76	0.75

a) ¿El diseño de la boquilla afecta el factor de la forma? Comparar las boquillas con un diagrama de dispersión y con un análisis de varianza, utilizando $\alpha = 0.05$.
b) Analizar los residuales de este experimento.
c) ¿Qué diseños de las boquillas son diferentes con respecto al factor de la forma? Trazar una gráfica del factor de la forma promedio para cada tipo de boquilla y compararla con una distribución t escalada. Comparar las conclusiones que se sacaron a partir de esta gráfica con las de la prueba del rango múltiple de Duncan.

4-6. Considere el experimento del algoritmo para controlar la proporción de alúmina del capítulo 3, sección 3-8. El experimento se llevó a cabo en realidad como un diseño de bloques aleatorizados, en el que se seleccionaron seis periodos como bloques, y se probaron los cuatro algoritmos para controlar la proporción en cada periodo. El voltaje promedio de la celda y la desviación estándar del voltaje (indicada entre paréntesis) para cada celda son los siguientes:

Algoritmo para controlar la proporción	Tiempo					
	1	2	3	4	5	6
1	4.93 (0.05)	4.86 (0.04)	4.75 (0.05)	4.95 (0.06)	4.79 (0.03)	4.88 (0.05)
2	4.85 (0.04)	4.91 (0.02)	4.79 (0.03)	4.85 (0.05)	4.75 (0.03)	4.85 (0.02)
3	4.83 (0.09)	4.88 (0.13)	4.90 (0.11)	4.75 (0.15)	4.82 (0.08)	4.90 (0.12)
4	4.89 (0.03)	4.77 (0.04)	4.94 (0.05)	4.86 (0.05)	4.79 (0.03)	4.76 (0.02)

a) Analizar los datos del voltaje promedio de las celdas. (Utilizar $\alpha = 0.05$.) ¿La elección del algoritmo para controlar la proporción afecta el voltaje promedio de las celdas?
b) Realizar el análisis apropiado de la desviación estándar del voltaje. (Recuerde que a éste se le llamó "ruido del crisol".) ¿La elección del algoritmo para controlar la proporción afecta el ruido del crisol?
c) Realizar los análisis residuales que parezcan apropiados.
d) ¿Qué algoritmo para controlar la proporción debería seleccionarse si el objetivo es reducir tanto el voltaje promedio de las celdas *como* el ruido del crisol?

4-7. El fabricante de una aleación maestra de aluminio produce refinadores de textura en forma de lingotes. La compañía produce el producto en cuatro hornos. Se sabe que cada horno tiene sus propias características únicas de operación, por lo que en cualquier experimento que se corra en la fundición en el que se use más de un horno, los hornos se considerarán como una variable perturbadora. Los ingenieros del proceso sospechan que la velocidad de agitación afecta la medida de la textura del producto. Cada horno puede operarse con

cuatro diferentes velocidades de agitación. Se lleva a cabo un diseño de bloques aleatorizados para un refinador particular y los datos resultantes de la medida de la textura se muestran a continuación:

	Horno			
Velocidad de agitación (rpm)	1	2	3	4
5	8	4	5	6
10	14	5	6	9
15	14	6	9	2
20	17	9	3	6

a) ¿Existe evidencia de que la velocidad de agitación afecta la medida de la textura?
b) Representar los residuales de este experimento en una gráfica de probabilidad normal. Interpretar esta gráfica.
c) Graficar los residuales contra el horno y la velocidad de agitación. ¿Esta gráfica proporciona alguna información útil?
d) ¿Cuál sería la recomendación de los ingenieros del proceso con respecto a la elección de la velocidad de agitación y del horno para este refinador de textura particular si es deseable una medida de la textura pequeña?

4-8. Analizar los datos del problema 4-2 utilizando la prueba general de significación de la regresión.

4-9. Suponiendo que los tipos de agentes químicos y los rollos de tela son fijos, estimar los parámetros del modelo τ_i y β_j del problema 4-1.

4-10. Trazar una curva de operación característica para el diseño del problema 4-2. ¿La prueba parece ser sensible a las diferencias pequeñas en los efectos de los tratamientos?

4-11. Suponga que falta la observación del agente químico 2 y el rollo 3 en el problema 4-1. Analizar el problema estimando el valor faltante. Realizar el análisis exacto y comparar los resultados.

4-12. ***Dos valores faltantes en un bloque aleatorizado.*** Suponga que en el problema 4-1 faltan las observaciones del agente químico tipo 2 y el rollo 3 y del agente químico tipo 4 y el rollo 4.
a) Analizar el diseño haciendo la estimación iterativa de los valores faltantes, como se describe en la sección 4-1.3.
b) Derivar SS_E con respecto a los dos valores faltantes, igualar los resultados con cero y resolver las ecuaciones para las estimaciones de los valores faltantes. Analizar el diseño utilizando estas dos estimaciones de los valores faltantes.
c) Deducir las fórmulas generales para estimar dos valores faltantes cuando las observaciones están en bloques *diferentes*.
d) Deducir las fórmulas generales para estimar dos valores faltantes cuando las observaciones están en el *mismo* bloque.

4-13. Un ingeniero industrial está realizando un experimento sobre el tiempo de enfoque del ojo. Se interesa en el efecto de la distancia del objeto al ojo sobre el tiempo de enfoque. Cuatro distancias diferentes son de interés. Cuenta con cinco sujetos para el experimento. Debido a que puede haber diferencias entre los individuos, el ingeniero decide realizar el experimento en un diseño de bloques aleatorizados. Los datos obtenidos se presentan a continuación. Analizar los datos de este experimento (utilizar $\alpha = 0.05$) y sacar las conclusiones apropiadas.

	Sujeto				
Distancia (pies)	1	2	3	4	5
4	10	6	6	6	6
6	7	6	6	1	6
8	5	3	3	2	5
10	6	4	4	2	3

4-14. Se estudia el efecto de cinco ingredientes diferentes (A, B, C, D y E) sobre el tiempo de reacción de un proceso químico. Cada lote de material nuevo sólo alcanza para permitir la realización de cinco corridas. Además, cada corrida requiere aproximadamente $1^1/_2$ horas, por lo que sólo pueden realizarse cinco corridas en un día. El experimentador decide realizar el experimento como un cuadrado latino para que los efectos del día y el lote puedan controlarse sistemáticamente. Obtiene los datos que se muestran enseguida. Analizar los datos de este experimento (utilizar $\alpha = 0.05$) y sacar conclusiones.

	Día				
Lote	1	2	3	4	5
1	$A = 8$	$B = 7$	$D = 1$	$C = 7$	$E = 3$
2	$C = 11$	$E = 2$	$A = 7$	$D = 3$	$B = 8$
3	$B = 4$	$A = 9$	$C = 10$	$E = 1$	$D = 5$
4	$D = 6$	$C = 8$	$E = 6$	$B = 6$	$A = 10$
5	$E = 4$	$D = 2$	$B = 3$	$A = 8$	$C = 8$

4-15. Un ingeniero industrial investiga el efecto de cuatro métodos de ensamblaje (A, B, C y D) sobre el tiempo de ensamblaje de un componente de televisores a color. Se seleccionan cuatro operadores para el estudio. Además, el ingeniero sabe que todos los métodos de ensamblaje producen fatiga, de tal modo que el tiempo requerido para el último ensamblaje puede ser mayor que para el primero, independientemente del método. Es decir, se desarrolla una tendencia en el tiempo de ensamblaje requerido. Para tomar en cuenta esta fuente de variabilidad, el ingeniero emplea el diseño del cuadrado latino que se presenta a continuación. Analizar los datos de este experimento ($\alpha = 0.05$) y sacar las conclusiones apropiadas.

Orden de	Operador			
ensamblaje	1	2	3	4
1	$C = 10$	$D = 14$	$A = 7$	$B = 8$
2	$B = 7$	$C = 18$	$D = 11$	$A = 8$
3	$A = 5$	$B = 10$	$C = 11$	$D = 9$
4	$D = 10$	$A = 10$	$B = 12$	$C = 14$

4-16. Suponga que en el problema 4-14 falta la observación del lote 3 en el día 4. Estimar el valor faltante con la ecuación 4-24, y realizar el análisis utilizando este valor.

4-17. Considere un cuadrado latino $p \times p$ con renglones (α_i), columnas (β_k) y tratamientos (τ_j) fijos. Obtener estimaciones de mínimos cuadrados de los parámetros del modelo α_i, β_k y τ_j.

4-18. Deducir la fórmula del valor faltante (ecuación 4-24) para el diseño del cuadrado latino.

4-19. ***Diseños que incluyen varios cuadrados latinos.*** (Ver Cochran y Cox [26] y John [61d].) El cuadrado latino $p \times p$ contiene únicamente p observaciones para cada tratamiento. Para obtener más réplicas, el experimentador puede usar varios cuadrados, por ejemplo n. No es relevante si los cuadrados usados son el mismo o son diferentes. El modelo apropiado es

$$y_{ijkh} = \mu + \rho_h + \alpha_{i(h)} + \tau_j + \beta_{k(h)} + (\tau\rho)_{jh} + \varepsilon_{ijkh} \begin{cases} i = 1, 2, \ldots, p \\ j = 1, 2, \ldots, p \\ k = 1, 2, \ldots, p \\ h = 1, 2, \ldots, n \end{cases}$$

donde y_{ijkh} es la observación del tratamiento j en el renglón i y la columna k del cuadrado h-ésimo. Observe que $\alpha_{i(h)}$ y $\beta_{k(h)}$ son los efectos del renglón y la columna en el cuadrado h-ésimo, ρ_h es el efecto del cuadrado h-ésimo y $(\tau\rho)_{jh}$ es la interacción entre los tratamientos y los cuadrados.

a) Establecer las ecuaciones normales para este modelo y resolverlas para las estimaciones de los parámetros del modelo. Suponga que las condiciones auxiliares apropiadas de los parámetros son $\Sigma_h\hat{\rho}_h = 0$, $\Sigma_i\hat{\alpha}_{i(h)} = 0$ y $\Sigma_k\hat{\beta}_{k(h)} = 0$ para cada h, $\Sigma_j\hat{\tau}_j = 0$, $\Sigma_j(\hat{\tau\rho})_{jh} = 0$ para cada h y $\Sigma_h(\hat{\tau\rho})_{jh} = 0$ para cada j.

b) Desarrollar la tabla del análisis de varianza para este diseño.

4-20. Comentar la forma en que pueden utilizarse las curvas de operación característica del apéndice con el diseño del cuadrado latino.

4-21. Suponga que en el problema 4-14 los datos tomados en el día 5 se analizaron incorrectamente y fue necesario descartarlos. Desarrollar un análisis apropiado para los datos restantes.

4-22. El rendimiento de un proceso químico se midió utilizando cinco lotes de materia prima, cinco concentraciones del ácido, cinco tiempos de procesamiento (*A, B, C, D* y *E*) y cinco concentraciones del catalizador ($\alpha, \beta, \gamma, \delta, \varepsilon$). Se usó el cuadrado grecolatino siguiente. Analizar los datos de este experimento (utilizar $\alpha = 0.05$) y sacar conclusiones.

	Concentración del ácido				
Lote	1	2	3	4	5
1	$A\alpha = 26$	$B\beta = 16$	$C\gamma = 19$	$D\delta = 16$	$E\varepsilon = 13$
2	$B\gamma = 18$	$C\delta = 21$	$D\varepsilon = 18$	$E\alpha = 11$	$A\beta = 21$
3	$C\varepsilon = 20$	$D\alpha = 12$	$E\beta = 16$	$A\gamma = 25$	$B\delta = 13$
4	$D\beta = 15$	$E\gamma = 15$	$A\delta = 22$	$B\varepsilon = 14$	$C\alpha = 17$
5	$E\delta = 10$	$A\varepsilon = 24$	$B\alpha = 17$	$C\beta = 17$	$D\gamma = 14$

4-23. Suponga que en el problema 4-15 el ingeniero sospecha que los sitios de trabajo usados por los cuatro operadores pueden representar una fuente adicional de variación. Es posible introducir un cuarto factor, el sitio de trabajo ($\alpha, \beta, \gamma, \delta$), y realizar otro experimento, de donde resulta el cuadrado grecolatino siguiente. Analizar los datos de este experimento (utilizar $\alpha = 0.05$) y sacar conclusiones.

Orden de	Operador			
ensamblaje	1	2	3	4
1	$C\beta = 11$	$B\gamma = 10$	$D\delta = 14$	$A\alpha = 8$
2	$B\alpha = 8$	$C\delta = 12$	$A\gamma = 10$	$D\beta = 12$
3	$A\delta = 9$	$D\alpha = 11$	$B\beta = 7$	$C\gamma = 15$
4	$D\gamma = 9$	$A\beta = 8$	$C\alpha = 18$	$B\delta = 6$

4-24. Construir un hipercuadrado 5 × 5 para estudiar los efectos de cinco factores. Desarrollar la tabla del análisis de varianza para este diseño.

4-25. Considere los datos de los problemas 4-15 y 4-23. Después de eliminar las letras griegas del problema 4-23, analizar los datos utilizando el método desarrollado en el problema 4-19.

4-26. Considere el diseño de bloques aleatorizados con un valor faltante en la tabla 4-7. Analizar los datos utilizando el análisis exacto del problema del valor faltante revisado en la sección 4-1.4. Comparar los resultados con el análisis aproximado de estos datos que se presenta en la tabla 4-8.

4-27. Un ingeniero estudia las características del rendimiento de combustible de cinco tipos de aditivos de gasolina. En la prueba de carretera el ingeniero desea usar los automóviles como bloques; sin embargo, debido a

una restricción de tiempo, debe utilizar un diseño de bloques incompletos. Realiza el diseño balanceado con los cinco bloques siguientes. Analizar los datos de este experimento (utilizar $\alpha = 0.05$) y sacar conclusiones.

	Automóvil				
Aditivo	1	2	3	4	5
1		17	14	13	12
2	14	14		13	10
3	12		13	12	9
4	13	11	11	12	
5	11	12	10		8

4-28. Construir un conjunto de contrastes ortogonales para los datos del problema 4-27. Calcular la suma de cuadrados para cada contraste.

4-29. Se estudian siete concentraciones diferentes de madera dura para determinar su efecto sobre la resistencia del papel producido. Sin embargo, en la planta piloto sólo pueden hacerse tres corridas de producción por día. Dado que los días pueden diferir, el analista utiliza el diseño de bloques incompletos balanceados que se muestra abajo. Analizar los datos de este experimento (utilizar $\alpha = 0.05$) y sacar conclusiones.

Concentración de	Días						
madera dura (%)	1	2	3	4	5	6	7
2	114				120		117
4	126	120				119	
6		137	117				134
8	141		129	149			
10		145		150	143		
12			120		118	123	
14				136		130	127

4-30. Analizar los datos del ejemplo 4-6 utilizando la prueba general de significación de la regresión.

4-31. Demostrar que $k\Sigma_{i=1}^{a}Q_i^2 / (\lambda a)$ es la suma de cuadrados ajustada de los tratamientos en un BIBD.

4-32. Un experimentador quiere comparar cuatro tratamientos en bloques de dos corridas. Encontrar un BIBD para este experimento con seis bloques.

4-33. Un experimentador quiere comparar ocho tratamientos en bloques de cuatro corridas. Encontrar un BIBD con 14 bloques y $\lambda = 3$.

4-34. Realizar el análisis interbloques del diseño del problema 4-27.

4-35. Realizar el análisis interbloques del diseño del problema 4-29.

4-36. Comprobar que no existe un BIBD con parámetros $a = 8$, $r = 8$, $k = 4$ y $b = 16$.

4-37. Demostrar que la varianza de los estimadores intrabloques $\{\hat{\tau}_i\}$ es $k(a-1)\sigma^2 / (\lambda a^2)$.

4-38. ***Diseños extendidos de bloques incompletos.*** Ocasionalmente, el tamaño del bloque cumple con la relación $a < k < 2a$. Un diseño extendido de bloques incompletos consiste en una sola réplica de cada tratamiento en cada bloque junto con un diseño de bloques incompletos con $k^* = k - a$. En el caso balanceado, el diseño de bloques incompletos tendrá los parámetros $k^* = k - a$, $r^* = r - b$ y λ^*. Desarrollar el análisis estadístico. (*Sugerencia:* en el diseño extendido de bloques incompletos, se tiene $\lambda = 2r - b + \lambda^*$.)

5 Introducción a los diseños factoriales

5-1 DEFINICIONES Y PRINCIPIOS BÁSICOS

En muchos experimentos interviene el estudio de los efectos de dos o más factores. En general, los **diseños factoriales** son los más eficientes para este tipo de experimentos. Por diseño factorial se entiende que en cada ensayo o réplica completa del experimento se investigan todas las combinaciones posibles de los niveles de los factores. Por ejemplo, si el factor A tiene a niveles y el factor B tiene b niveles, cada réplica contiene todas las ab combinaciones de los tratamientos. Cuando los factores están incluidos en un diseño factorial, es común decir que están **cruzados**.

El efecto de un factor se define como el cambio en la respuesta producido por un cambio en el nivel del factor. Con frecuencia se le llama **efecto principal** porque se refiere a los factores de interés primario en el experimento. Por ejemplo, considere el experimento sencillo de la figura 5-1. Se trata de un experimento factorial de dos factores en el que los dos factores del diseño tienen dos niveles. A estos niveles se les ha denominado "bajo" y "alto" y se denotan como "–" y "+", respectivamente. El efecto principal del factor A de este diseño de dos niveles puede visualizarse como la diferencia entre la respuesta promedio con el nivel bajo de A y la respuesta promedio con el nivel alto de A. Numéricamente, esto es

$$A = \frac{40+52}{2} - \frac{20+30}{2} = 21$$

Es decir, cuando el factor A se incrementa del nivel bajo al nivel alto se produce un **incremento de la respuesta promedio** de 21 unidades. De manera similar, el efecto principal de B es

$$B = \frac{30+52}{2} - \frac{20+40}{2} = 11$$

Cuando los factores tienen más de dos niveles, es necesario modificar el procedimiento anterior, ya que existen otras formas de definir el efecto de un factor. Este punto se estudia con mayor profundidad más adelante.

En algunos experimentos puede encontrarse que la diferencia en la respuesta entre los niveles de un factor no es la misma para todos los niveles de los otros factores. Cuando esto ocurre, existe una **interac-**

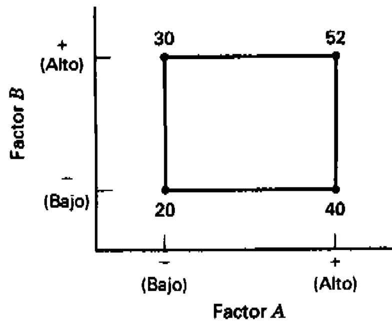


Figura 5-1 Experimento factorial de dos factores con la respuesta (y) indicada en los vértices.

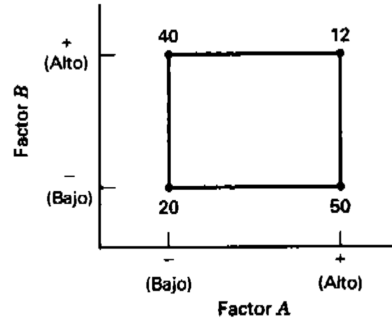


Figura 5-2 Experimento factorial de dos factores con interacción.

ción entre los factores. Por ejemplo, considere el experimento factorial de dos factores que se ilustra en la figura 5-2. Con el nivel bajo del factor B (o B^-), el efecto de A es

$$A = 50 - 20 = 30$$

y con el nivel alto del factor B (o B^+), el efecto de A es

$$A = 12 - 40 = -28$$

Puesto que el efecto de A depende del nivel que se elige para el factor B, se observa que existe una interacción entre A y B. La magnitud del efecto de la interacción es la *diferencia* promedio de estos dos efectos de A, o $AB = (-28 - 30)/2 = -29$. Evidentemente, en este experimento la interacción es grande.

Estas ideas pueden ilustrarse gráficamente. En la figura 5-3 se grafican los datos de las respuestas de la figura 5-1 contra el factor A para ambos niveles del factor B. Observe que las rectas B^- y B^+ son aproximadamente paralelas, lo cual indica la ausencia de interacción entre los factores A y B. De manera similar, en la figura 5-4 se grafican los datos de las respuestas de la figura 5-2. En este caso se observa que las rectas B^- y B^+ no son paralelas. Esto indica una interacción entre los factores A y B. Gráficas como éstas son de gran ayuda para interpretar las interacciones significativas y para reportar los resultados al personal sin preparación estadística. Sin embargo, no deberán utilizarse como la única técnica para el análisis de datos, ya que su interpretación es subjetiva y su apariencia con frecuencia es engañosa.

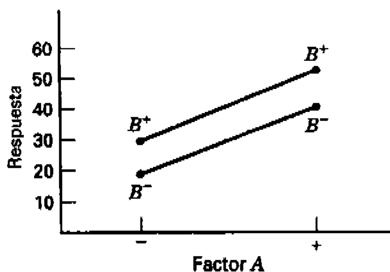


Figura 5-3 Experimento factorial sin interacción.

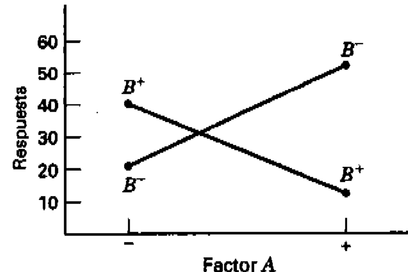


Figura 5-4 Experimento factorial con interacción.

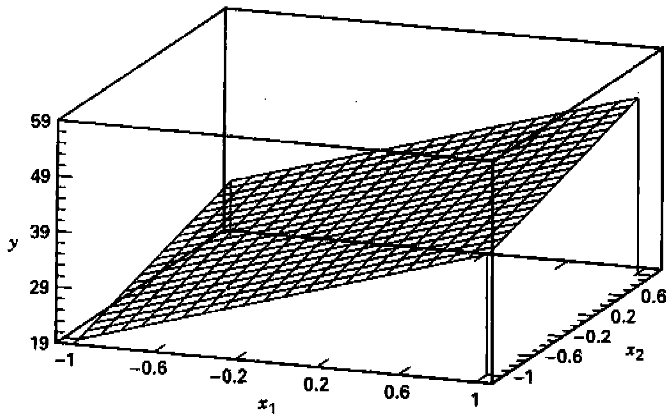
El concepto de interacción puede ilustrarse de otra manera. Suponga que los dos factores del diseño tratado son **cuantitativos** (temperatura, presión, tiempo, etc.). Entonces una **representación con un modelo de regresión** del experimento factorial de dos factores podría escribirse como

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \varepsilon$$

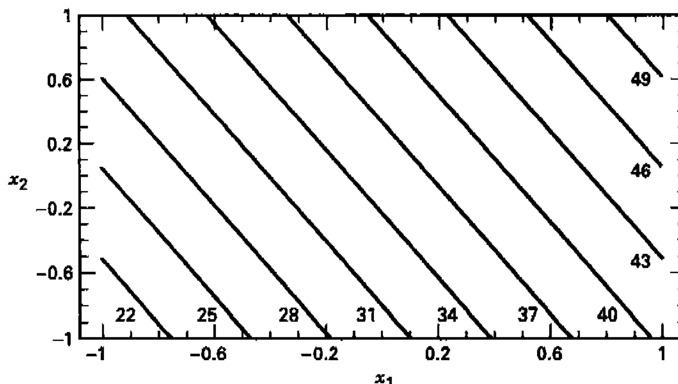
donde y es la respuesta, las β son parámetros cuyos valores deben determinarse, x_1 es una variable que representa al factor A, x_2 es una variable que representa al factor B, y ε es un término del error aleatorio. Las variables x_1 y x_2 se definen en una **escala codificada** de -1 a $+1$ (los niveles bajo y alto de A y B), y x_1x_2 representa la interacción entre x_1 y x_2.

Las estimaciones de los parámetros en este modelo de regresión resultan estar relacionadas con las estimaciones de los efectos. Para el experimento ilustrado en la figura 5-1 se encuentra que los efectos principales de A y B son $A = 21$ y $B = 11$. Las estimaciones de β_1 y β_2 son la mitad del valor del efecto principal correspondiente; por lo tanto, $\hat{\beta}_1 = 21/2 = 10.5$ y $\hat{\beta}_2 = 11/2 = 5.5$. El efecto de la interacción de la figura 5-1 es $AB = 1$, por lo que el valor del coeficiente de la interacción en el modelo de regresión es $\hat{\beta}_{12} = 1/2 = 0.5$. El parámetro β_0 se estima con el promedio de las cuatro respuestas, o $\hat{\beta}_0 = (20+40+30+52)/4 = 35.5$. Por lo tanto, el modelo de regresión ajustado es

$$\hat{y} = 35.5 + 10.5x_1 + 5.5x_2 + 0.5x_1x_2$$



a) La superficie de respuesta



b) La gráfica de contorno

Figura 5-5 La superficie de respuesta y la gráfica de contorno para el modelo $\hat{y} = 35.5 + 10.5x_1 + 5.5x_2$.

Las estimaciones de los parámetros obtenidas de esta manera para el diseño factorial en el que todos los factores tienen dos niveles (– y +) resultan ser **estimaciones de mínimos cuadrados** (se abundará sobre el tema más adelante).

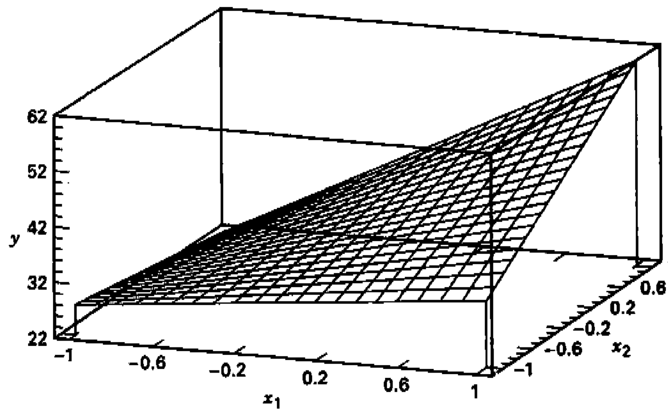
El coeficiente de la interacción ($\hat{\beta}_{12} = 0.5$) es pequeño en comparación con los coeficientes de los efectos principales $\hat{\beta}_1$ y $\hat{\beta}_2$. La interpretación que se hará de este hecho es que la interacción es pequeña y puede ignorarse. Por lo tanto, al eliminar el término $0.5x_1x_2$ se obtiene el modelo

$$\hat{y} = 35.5 + 10.5x_1 + 5.5x_2$$

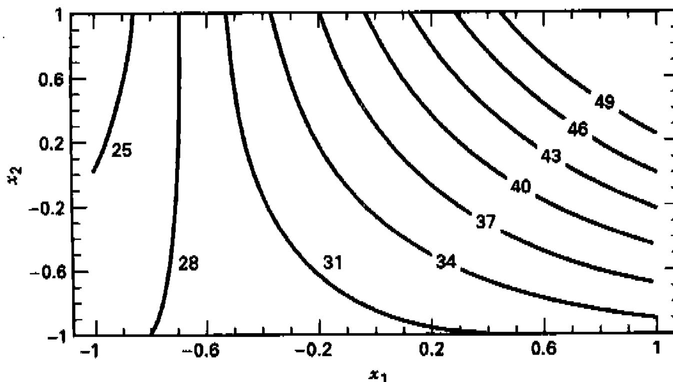
En la figura 5-5 se muestran las representaciones gráficas de este modelo. En la figura 5-5*a* se tiene una gráfica del plano de los valores de y generados por las diferentes combinaciones de x_1 y x_2. A esta gráfica tridimensional se le llama **gráfica de superficie de respuesta.** En la figura 5-5*b* se muestran las líneas de contorno para las respuestas constantes y en el plano x_1, x_2. Observe que como la superficie de respuesta es un plano, la gráfica de contorno contiene líneas rectas paralelas.

Suponga ahora que la contribución de la interacción en el experimento no fuera insignificante; es decir, que el coeficiente β_{12} no fuera pequeño. En la figura 5-6 se presenta la superficie de respuesta y la gráfica de contorno del modelo

$$\hat{y} = 35.5 + 10.5x_1 + 5.5x_2 + 8x_1x_2$$



a) La superficie de respuesta



b) La gráfica de contorno

Figura 5-6 La superficie de respuesta y la gráfica de contorno para el modelo $\hat{y} = 35.5 + 10.5x_1 + 5.5x_2 + 8x_1x_2$.

(Se ha hecho que el efecto de la interacción sea el promedio de los dos efectos principales.) Observe que el efecto significativo de la interacción provoca el "torcimiento" del plano de la figura 5-6*a*. Este torcimiento de la superficie de respuesta produce líneas de contorno curvas para las respuestas constantes en el plano x_1, x_2, como se muestra en la figura 5-6*b*. Por lo tanto, una **interacción es una forma de curvatura** en el modelo de superficie de respuesta fundamental del experimento.

El modelo de superficie de respuesta de un experimento es de gran importancia y utilidad. El tema se ampliará en la sección 5-5 y en capítulos posteriores.

En general, cuando una interacción es grande, los efectos principales correspondientes tienen escaso significado práctico. En el experimento de la figura 5-2, la estimación del efecto principal de A sería

$$A = \frac{50+12}{2} - \frac{20+40}{2} = 1$$

que es muy pequeño, y se llegaría a concluir que no hay ningún efecto debido a A. Sin embargo, cuando se examinan los efectos de A con *niveles diferentes del factor B*, se observa que no es éste el caso. El factor A tiene un efecto, pero *depende del nivel del factor B*. Es decir, el conocimiento de la interacción AB es más útil que el conocimiento del efecto principal. Una interacción significativa suele **enmascarar** la significación de los efectos principales. Estos puntos se ponen de manifiesto con claridad en la gráfica de la interacción de la figura 5-4. En presencia de una interacción significativa, el experimentador deberá por lo general examinar los niveles de uno de los factores, por ejemplo del factor A, manteniendo fijos los niveles de los otros factores para sacar conclusiones acerca del efecto principal de A.

5-2 LA VENTAJA DE LOS DISEÑOS FACTORIALES

Es sencillo ilustrar la ventaja de los diseños factoriales. Suponga que se tienen dos factores A y B, cada uno con dos niveles. Los niveles de los factores se denotan por A^-, A^+, B^- y B^+. Podría obtenerse información acerca de ambos factores haciéndolos variar uno a la vez, como se muestra en la figura 5-7. El efecto de cambiar el factor A está dado por $A^+B^- - A^-B^-$, y el efecto de cambiar el factor B está dado por $A^-B^+ - A^-B^-$. Debido a que está presente el error experimental, es deseable realizar dos observaciones, por ejemplo, para cada combinación de tratamientos y estimar los efectos de los factores utilizando las respuestas promedio. Por lo tanto, se necesita un total de seis observaciones.

Si se hubiera efectuado un experimento factorial, se habría registrado una combinación adicional de los tratamientos, A^+B^+. Ahora, utilizando sólo *cuatro* observaciones, pueden hacerse dos estimaciones del efecto de A: $A^+B^- - A^-B^-$ y $A^+B^+ - A^-B^+$. De manera similar, pueden hacerse dos estimaciones del

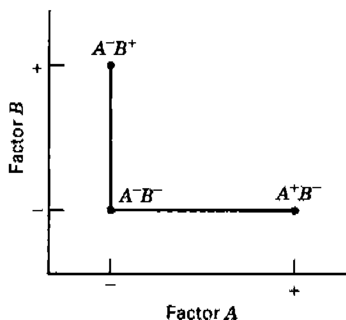


Figura 5-7 Experimento con un factor a la vez.

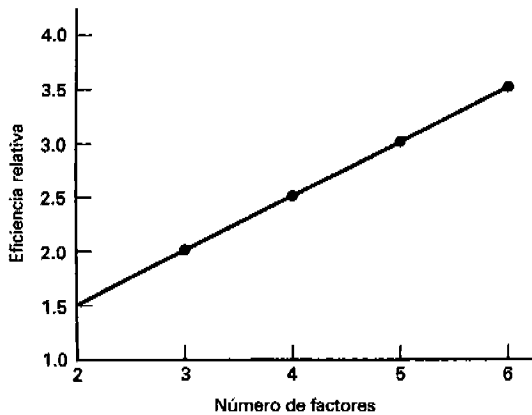


Figura 5-8 Eficiencia relativa de un diseño factorial con respecto a un experimento de un factor a la vez (dos niveles del factor).

efecto de B. Estas dos estimaciones de cada efecto principal podrían promediarse para producir efectos principales promedio que tienen la *misma precisión* que las estimaciones del experimento con un solo factor, pero sólo se requieren cuatro observaciones en total, y nosotros diríamos que la eficiencia relativa del diseño factorial con respecto al experimento de un factor a la vez es de (6/4) = 1.5. En general, esta eficiencia relativa aumentará conforme se incremente el número de factores, como se muestra en la figura 5-8.

Suponga ahora que está presente una interacción. Si el diseño de un factor a la vez indicara que A^-B^+ y A^+B^- dieron mejores respuestas que A^-B^-, una conclusión lógica sería que A^+B^+ sería todavía mejor. Sin embargo, si está presente una interacción, esta conclusión puede ser una *equivocación grave*. Para un ejemplo, referirse al experimento de la figura 5-2.

En resumen, observe que los diseños factoriales ofrecen varias ventajas. Son más eficientes que los experimentos de un factor a la vez. Además, un diseño factorial es necesario cuando puede haber interacciones presentes a fin de evitar llegar a conclusiones incorrectas. Por último, los diseños factoriales permiten la estimación de los efectos de un factor con varios niveles de los factores restantes, produciendo conclusiones que son válidas para un rango de condiciones experimentales.

5-3 DISEÑO FACTORIAL DE DOS FACTORES

5-3.1 Un ejemplo

Los tipos más simples de diseños factoriales incluyen únicamente dos factores o conjuntos de tratamientos. Hay a niveles del factor A y b niveles del factor B, los cuales se disponen en un diseño factorial; es decir, cada réplica del experimento contiene todas las ab combinaciones de los tratamientos. En general, hay n réplicas.

Como ejemplo de un diseño factorial en el que intervienen dos factores, un ingeniero está diseñando una batería que se usará en un dispositivo que se someterá a variaciones de temperatura extremas. El único parámetro del diseño que puede seleccionar en este punto es el material de la placa o ánodo de la batería, y tiene tres elecciones posibles. Cuando el dispositivo esté fabricado y se envíe al campo, el ingeniero no tendrá control sobre las temperaturas extremas en las que operará el dispositivo, pero sabe por expe-

Tabla 5-1 Datos de la vida (en horas) para el ejemplo del diseño de la batería

Tipo de material	Temperatura (°F)					
	15		70		125	
1	130	155	34	40	20	70
	74	180	80	75	82	58
2	150	188	136	122	25	70
	159	126	106	115	58	45
3	138	110	174	120	96	104
	168	160	150	139	82	60

riencia que la temperatura probablemente afectará la vida efectiva de la batería. Sin embargo, la temperatura puede controlarse en el laboratorio donde se desarrolla el producto para fines de prueba.

El ingeniero decide probar los tres materiales de la placa con tres niveles de temperatura —15, 70 y 125°F—, ya que estos niveles de temperatura son consistentes con el medio ambiente donde se usará finalmente el producto. Se prueban cuatro baterías con cada combinación del material de la placa y la temperatura, y las 36 pruebas se corren de manera aleatoria. En la tabla 5-1 se presentan los datos del experimento y de la vida observada de la batería.

En este problema, el ingeniero quiere responder las preguntas siguientes:

1. ¿Qué efectos tienen el tipo de material y la temperatura sobre la vida de la batería?
2. ¿Existe alguna elección del material que produzca *de manera regular una vida larga* de la batería *independientemente de la temperatura*?

La segunda pregunta es de particular importancia. Quizá sea posible encontrar una alternativa del material que no resulte afectada considerablemente por la temperatura. De ser éste el caso, el ingeniero puede hacer que la batería sea **robusta** para la variación de la temperatura en el campo. Se trata de un ejemplo de la aplicación del diseño experimental estadístico en el **diseño de productos robustos,** un problema de ingeniería muy importante.

El anterior es un ejemplo específico del caso general de un diseño factorial de dos factores. Para pasar al caso general, sea y_{ijk} la respuesta observada cuando el factor A tiene el nivel i-ésimo ($i = 1, 2, ..., a$) y el factor B tiene el nivel j-ésimo ($j = 1, 2, ..., b$) en la réplica k-ésima ($k = 1, 2, ..., n$). En general, el experimento factorial de dos factores aparecerá como en la tabla 5-2. El orden en que se hacen las abn observaciones se selecciona al azar, por lo que este diseño es un **diseño completamente aleatorizado.**

Tabla 5-2 Arreglo general de un diseño factorial de dos factores

		Factor B			
		1	2	...	b
Factor A	1	$y_{111}, y_{112}, \ldots, y_{11n}$	$y_{121}, y_{122}, \ldots, y_{12n}$		$y_{1b1}, y_{1b2}, \ldots, y_{1bn}$
	2	$y_{211}, y_{212}, \ldots, y_{21n}$	$y_{221}, y_{222}, \ldots, y_{22n}$		$y_{2b1}, y_{2b2}, \ldots, y_{2bn}$
	⋮				
	a	$y_{a11}, y_{a12}, \ldots, y_{a1n}$	$y_{a21}, y_{a22}, \ldots, y_{a2n}$		$y_{ab1}, y_{ab2}, \ldots, y_{abn}$

Las observaciones de un experimento factorial pueden describirse con un modelo. Hay varias formas de escribir el modelo de un experimento factorial. El **modelo de los efectos** es

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \varepsilon_{ijk} \quad \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, b \\ k = 1, 2, \ldots, n \end{cases} \tag{5-1}$$

donde μ es el efecto promedio global, τ_i es el efecto del nivel i-ésimo del factor A de los renglones, β_j es el efecto del nivel j-ésimo del factor B de las columnas, $(\tau\beta)_{ij}$ es el efecto de la interacción entre τ_i y β_j, y ε_{ijk} es un componente del error aleatorio. Se supone que ambos factores son **fijos**, y los efectos de los tratamientos se definen como las desviaciones de la media global, por lo que $\Sigma_{i=1}^{a}\tau_i = 0$ y $\Sigma_{j=1}^{b}\beta_j = 0$. De manera similar, los efectos de las interacciones son fijos y se definen de tal modo que $\Sigma_{i=1}^{a}(\tau\beta)_{ij} = \Sigma_{j=1}^{b}(\tau\beta)_{ij} = 0$. Puesto que hay n réplicas del experimento, hay abn observaciones en total.

Otro modelo posible de un experimento factorial es el **modelo de las medias**

$$y_{ijk} = \mu_{ij} + \varepsilon_{ijk} \quad \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, b \\ k = 1, 2, \ldots, n \end{cases}$$

donde la media de la celda ij-ésima es

$$\mu_{ij} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij}$$

También podría usarse un **modelo de regresión** como en la sección 5-1. Los modelos de regresión resultan particularmente útiles cuando uno o más de los factores del experimento son cuantitativos. En la mayor parte de este capítulo se usará el modelo de los efectos (ecuación 5-1) con referencia al modelo de regresión en la sección 5-5.

En el diseño factorial de dos factores, los factores (o tratamientos) de los renglones y las columnas, A y B, son de igual interés. Específicamente, el interés se encuentra en **probar hipótesis** acerca de la igualdad de los efectos de los tratamientos de los renglones, por ejemplo,

$$\begin{aligned} &H_0: \tau_1 = \tau_2 = \cdots = \tau_a = 0 \\ &H_1: \text{al menos una } \tau_i \neq 0 \end{aligned} \tag{5-2a}$$

y de la igualdad de los efectos de los tratamientos de las columnas, por ejemplo,

$$\begin{aligned} &H_0: \beta_1 = \beta_2 = \cdots = \beta_b = 0 \\ &H_1: \text{al menos una } \beta_j \neq 0 \end{aligned} \tag{5-2b}$$

También existe interés en determinar si los tratamientos de los renglones y las columnas *interactúan*. Por lo tanto, también querría probarse

$$\begin{aligned} &H_0: (\tau\beta)_{ij} = 0 \qquad \text{para todas las } i,\ j \\ &H_1: \text{al menos una } (\tau\beta)_{ij} \neq 0 \end{aligned} \tag{5-2c}$$

A continuación se indica cómo se prueban estas hipótesis utilizando un **análisis de varianza de dos factores.**

5-3.2 Análisis estadístico del modelo con efectos fijos

Sea que $y_{i..}$ denote el total de observaciones bajo el nivel i-ésimo del factor A, que $y_{.j.}$ denote el total de observaciones bajo el nivel j-ésimo del factor B, que $y_{ij.}$ denote el total de observaciones de la celda

ij-ésima, y que $y_{...}$ denote el gran total de todas las observaciones. Se definen $\bar{y}_{i..}, \bar{y}_{.j.}, \bar{y}_{ij.}$ y $\bar{y}_{...}$ como los promedios correspondientes de los renglones, las columnas, las celdas y el gran promedio. Expresado matemáticamente,

$$
\begin{aligned}
y_{i..} &= \sum_{j=1}^{b}\sum_{k=1}^{n} y_{ijk} & \bar{y}_{i..} &= \frac{y_{i..}}{bn} & i &= 1, 2,\ldots, a \\
y_{.j.} &= \sum_{i=1}^{a}\sum_{k=1}^{n} y_{ijk} & \bar{y}_{.j.} &= \frac{y_{.j.}}{an} & j &= 1, 2,\ldots, b \\
y_{ij.} &= \sum_{k=1}^{n} y_{ijk} & \bar{y}_{ij.} &= \frac{y_{ij.}}{n} & i &= 1, 2,\ldots, a \quad j = 1, 2,\ldots, b \\
y_{...} &= \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} y_{ijk} & \bar{y}_{...} &= \frac{y_{...}}{abn}
\end{aligned}
\qquad (5\text{-}3)
$$

La **suma de cuadrados total corregida** puede escribirse como

$$
\begin{aligned}
\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} (y_{ijk} - \bar{y}_{...})^2 &= \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} [(\bar{y}_{i..} - \bar{y}_{...}) + (\bar{y}_{.j.} - \bar{y}_{...}) \\
&\quad + (\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...}) + (y_{ijk} - \bar{y}_{ij.})]^2 \\
&= bn\sum_{i=1}^{a} (\bar{y}_{i..} - \bar{y}_{...})^2 + an\sum_{j=1}^{b} (\bar{y}_{.j.} - \bar{y}_{...})^2 \\
&\quad + n\sum_{i=1}^{a}\sum_{j=1}^{b} (\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...})^2 \\
&\quad + \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} (y_{ijk} - \bar{y}_{ij.})^2
\end{aligned}
\qquad (5\text{-}4)
$$

ya que los seis productos cruzados del lado derecho de la igualdad son cero. Observe que se ha hecho la partición de la suma de cuadrados total en una suma de cuadrados debida a "los renglones", o factor A (SS_A); una suma de cuadrados debida a "las columnas", o factor B (SS_B); una suma de cuadrados debida a la interacción entre A y B (SS_{AB}); y una suma de cuadrados debida al error (SS_E). Por el último componente del lado derecho de la igualdad de la ecuación 5-4, se observa que debe haber por lo menos dos réplicas ($n \geq 2$) para obtener una suma de cuadrados del error.

La ecuación 5-4 puede escribirse simbólicamente como

$$SS_T = SS_A + SS_B + SS_{AB} + SS_E \qquad (5\text{-}5)$$

El número de grados de libertad asociado con cada suma de cuadrados es

Efecto	Grados de libertad
A	$a - 1$
B	$b - 1$
Interacción AB	$(a - 1)(b - 1)$
Error	$ab(n - 1)$
Total	$abn - 1$

Esta asignación de los $abn - 1$ grados de libertad totales a las sumas de cuadrados puede justificarse de la siguiente manera: los efectos principales A y B tienen a y b niveles, respectivamente; por lo tanto, tienen $a - 1$ y $b - 1$ grados de libertad, como se indica. Los grados de libertad de la interacción son sólo el número de grados de libertad de las celdas (que es $ab - 1$) menos el número de grados de libertad de los dos efectos principales A y B; es decir, $ab - 1 - (a - 1) - (b - 1) = (a - 1)(b - 1)$. Dentro de cada una de las ab celdas hay $n - 1$ grados de libertad entre las n réplicas; por lo tanto hay $ab(n - 1)$ grados de libertad para el error. Observe que la suma del número de grados de libertad en el lado derecho de la ecuación 5-5 es igual al número total de grados de libertad.

Cada suma de cuadrados dividida por sus grados de libertad es un **cuadrado medio.** Los valores esperados de los cuadrados medios son

$$E(MS_A) = E\left(\frac{SS_A}{a-1}\right) = \sigma^2 + \frac{bn\sum_{i=1}^{a} \tau_i^2}{a-1}$$

$$E(MS_B) = E\left(\frac{SS_B}{b-1}\right) = \sigma^2 + \frac{an\sum_{j=1}^{b} \beta_j^2}{b-1}$$

$$E(MS_{AB}) = E\left(\frac{SS_{AB}}{(a-1)(b-1)}\right) = \sigma^2 + \frac{n\sum_{i=1}^{a}\sum_{j=1}^{b} (\tau\beta)_{ij}^2}{(a-1)(b-1)}$$

y

$$E(MS_E) = E\left(\frac{SS_E}{ab(n-1)}\right) = \sigma^2$$

Observe que si es verdadera la hipótesis nula de que no hay efectos de los tratamientos de los renglones, ni de los tratamientos de las columnas, ni interacción, entonces MS_A, MS_B, MS_{AB} y MS_E son todas estimaciones de σ^2. Sin embargo, si hay diferencias entre los efectos de los tratamientos de los renglones, por ejemplo, entonces MS_A será mayor que MS_E. De manera similar, si están presentes efectos de los tratamientos de las columnas o de la interacción, entonces los cuadrados medios correspondientes serán mayores que MS_E. Por lo tanto, para probar la significación de los dos efectos principales y su interacción, simplemente se divide el cuadrado medio correspondiente por el cuadrado medio del error. Los valores grandes de este cociente implican que los datos no apoyan la hipótesis nula.

Si se supone que el modelo (ecuación 5-1) es adecuado y que los términos del error ε_{ijk} tienen una distribución normal e independiente con varianza σ^2 constante, entonces cada uno de los cocientes de cuadrados medios MS_A/MS_E, MS_B/MS_E y MS_{AB}/MS_E se distribuyen como F con $a - 1$, $b - 1$ y $(a - 1)(b - 1)$ grados de libertad en el numerador, respectivamente, y $ab(n - 1)$ grados de libertad en el denominador,[1] y la región crítica sería la cola superior de la distribución F. El procedimiento de prueba suele resumirse en una **tabla del análisis de varianza,** como se muestra en la tabla 5-3.

En lo que a los cálculos se refiere, por lo general se emplea un paquete de software de estadística para realizar el análisis de varianza. Sin embargo, no es complicado obtener fórmulas para calcular manual-

[1] La prueba F puede considerarse como una aproximación de una prueba de aleatorización, como se señaló anteriormente.

Tabla 5-3 La tabla del análisis de varianza para el diseño factorial de dos factores, modelo con efectos fijos

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0
Tratamientos A	SS_A	$a-1$	$MS_A = \frac{SS_A}{a-1}$	$F_0 = \frac{MS_A}{MS_E}$
Tratamientos B	SS_B	$b-1$	$MS_B = \frac{SS_B}{b-1}$	$F_0 = \frac{MS_B}{MS_E}$
Interacción	SS_{AB}	$(a-1)(b-1)$	$MS_{AB} = \frac{SS_{AB}}{(a-1)(b-1)}$	$F_0 = \frac{MS_{AB}}{MS_E}$
Error	SS_E	$ab(n-1)$	$MS_E = \frac{SS_E}{ab(n-1)}$	
Total	SS_T	$abn-1$		

mente las sumas de cuadrados de la ecuación 5-5. La suma de cuadrados total se calcula como de costumbre con

$$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} y_{ijk}^2 - \frac{y_{...}^2}{abn} \tag{5-6}$$

Las sumas de cuadrados de los efectos principales son

$$SS_A = \frac{1}{bn}\sum_{i=1}^{a} y_{i..}^2 - \frac{y_{...}^2}{abn} \tag{5-7}$$

y

$$SS_B = \frac{1}{an}\sum_{j=1}^{b} y_{.j.}^2 - \frac{y_{...}^2}{abn} \tag{5-8}$$

Es conveniente obtener SS_{AB} en dos pasos. Se calcula primero la suma de cuadrados entre los totales de las ab celdas, a la que se denomina la suma de cuadrados debida a los "subtotales":

$$SS_{\text{Subtotales}} = \frac{1}{n}\sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij.}^2 - \frac{y_{...}^2}{abn}$$

Esta suma de cuadrados también contiene a SS_A y SS_B. Por lo tanto, el segundo paso consiste en calcular SS_{AB} como

$$SS_{AB} = SS_{\text{Subtotales}} - SS_A - SS_B \tag{5-9}$$

Puede calcularse SS_E por sustracción como

$$SS_E = SS_T - SS_{AB} - SS_A - SS_B \tag{5-10}$$

o

$$SS_E = SS_T - SS_{\text{Subtotales}}$$

EJEMPLO 5-1 ··

El experimento del diseño de la batería

En la tabla 5-4 se presenta la vida efectiva (en horas) observada en el ejemplo del diseño de la batería que se describió en la sección 5-3.1. Los totales de los renglones y las columnas se indican en los márgenes de la tabla y los números encerrados en un círculo son los totales de las celdas.

Tabla 5-4 Datos de la vida (en horas) del experimento del diseño de la batería

Tipo de material	Temperatura (°F) 15			70			125			$y_{i..}$
1	130 74	155 180	(539)	34 80	40 75	(229)	20 82	70 58	(230)	998
2	150 159	188 126	(623)	136 106	122 115	(479)	25 58	70 45	(198)	1300
3	138 168	110 160	(576)	174 150	120 139	(583)	96 82	104 60	(342)	1501
$y_{.j.}$	1738			1291			770			$3799 = y_{...}$

Las sumas de cuadrados se calculan de la siguiente manera:

$$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} y_{ijk}^2 - \frac{y_{...}^2}{abn}$$

$$= (130)^2 + (155)^2 + (74)^2 + \cdots + (60)^2 - \frac{(3799)^2}{36} = 77{,}646.97$$

$$SS_{\text{Material}} = \frac{1}{bn}\sum_{i=1}^{a} y_{i..}^2 - \frac{y_{...}^2}{abn}$$

$$= \frac{1}{(3)(4)}[(998)^2 + (1300)^2 + (1501)^2] - \frac{(3799)^2}{36} = 10{,}683.72$$

$$SS_{\text{Temperatura}} = \frac{1}{an}\sum_{j=1}^{b} y_{.j.}^2 - \frac{y_{...}^2}{abn}$$

$$= \frac{1}{(3)(4)}[(1738)^2 + (1291)^2 + (770)^2] - \frac{(3799)^2}{36} = 39{,}118.72$$

$$SS_{\text{Interacción}} = \frac{1}{n}\sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij.}^2 - \frac{y_{...}^2}{abn} - SS_{\text{Material}} - SS_{\text{Temperatura}}$$

$$= \frac{1}{4}[(539)^2 + (229)^2 + \cdots + (342)^2] - \frac{(3799)^2}{36} - 10{,}683.72$$
$$- 39{,}118.72 = 9613.78$$

y

$$SS_E = SS_T - SS_{\text{Material}} - SS_{\text{Temperatura}} - SS_{\text{Interacción}}$$
$$= 77{,}646.97 - 10{,}683.72 - 39{,}118.72 - 9613.78 = 18{,}230.75$$

En la tabla 5-5 se muestra el análisis de varianza. Puesto que $F_{0.05,4,27} = 2.73$, se concluye que hay una interacción significativa entre los tipos del material y la temperatura. Además, $F_{0.05,2,27} = 3.35$, por lo que los efectos principales del tipo de material y la temperatura también son significativos. En la tabla 5-5 también se muestran los valores P para los estadísticos de la prueba.

Como ayuda para interpretar los resultados de este experimento, es conveniente construir una gráfica de las respuestas promedio para cada combinación de los tratamientos. Esta gráfica se muestra en la figura 5-9. El hecho de que las rectas no sean paralelas indica que la interacción es significativa. En general, se consigue una vida más larga con una temperatura baja, independientemente del tipo de material. Al cambiar de una temperatura baja a una intermedia, la vida de la batería con el material tipo 3 *tiene un in-*

Tabla 5-5 Análisis de varianza de los datos de la vida de la batería

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0	Valor P
Tipos de material	10,683.72	2	5,341.86	7.91	0.0020
Temperatura	39,118.72	2	19,559.36	28.97	0.0001
Interacción	9,613.78	4	2,403.44	3.56	0.0186
Error	18,230.75	27	675.21		
Total	77,646.97	35			

cremento real, mientras que con los materiales tipos 1 y 2 disminuye. Con una temperatura de intermedia a alta, la vida de la batería disminuye para los materiales tipos 2 y 3 y se mantiene en esencia sin cambio para el material tipo 1. El material tipo 3 parece producir los mejores resultados si se quiere una pérdida menor de la vida efectiva cuando la temperatura cambia.

Comparaciones múltiples

Cuando el análisis de varianza indica que las medias de los renglones o las columnas difieren, por lo general es de interés hacer comparaciones entre las medias individuales de los renglones o las columnas para descubrir diferencias específicas. Los métodos de comparaciones múltiples revisados en el capítulo 3 son útiles a este respecto.

Se ilustra ahora el uso de la prueba de Tukey con los datos de la vida de la batería del ejemplo 5-1. Observe que en este experimento, la interacción es significativa. Cuando la interacción es significativa, las comparaciones entre las medias de uno de los factores (por ejemplo, A) pueden ser oscurecidas por la interacción AB. Una forma de abordar esta cuestión consiste en fijar el factor B en un nivel específico y aplicar la prueba de Tukey a las medias del factor A con ese nivel. Para ilustrar, suponga que en el ejemplo 5-1 el interés se encuentra en detectar las diferencias entre las medias de los tres tipos de material. Puesto que la interacción es significativa, esta comparación se hace con un solo nivel de la temperatura, por ejemplo el nivel 2 (70°F). Se supone que la mejor estimación de la varianza del error es MS_E de la tabla del análisis de varianza, utilizando el supuesto de que la varianza del error experimental es la misma para todas las combinaciones de tratamientos.

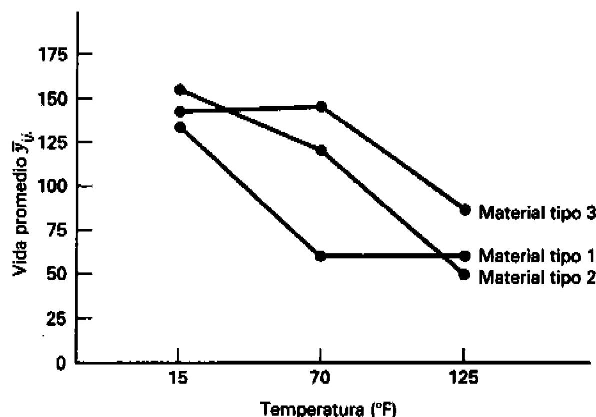


Figura 5-9 Gráfica tipo de material-temperatura para el ejemplo 5-1.

Los promedios de los tres tipos de material a 70°F dispuestos en orden ascendente son

$$\bar{y}_{12.} = 57.25 \quad \text{(material tipo 1)}$$
$$\bar{y}_{22.} = 119.75 \quad \text{(material tipo 2)}$$
$$\bar{y}_{32.} = 145.75 \quad \text{(material tipo 3)}$$

y

$$\begin{aligned} T_{0.05} &= q_{0.05}(3, 27)\sqrt{\frac{MS_E}{n}} \\ &= 3.50\sqrt{\frac{675.21}{4}} \\ &= 45.47 \end{aligned}$$

donde $q_{0.05}(3, 27) = 3.50$ se obtiene por interpolación en la tabla VIII del apéndice. Las comparaciones por pares dan como resultado

$$\begin{aligned} &3 \text{ vs. } 1: \quad 145.75 - 57.25 = 88.50 > T_{0.05} = 45.47 \\ &3 \text{ vs. } 2: \quad 145.75 - 119.75 = 26.00 < T_{0.05} = 45.47 \\ &2 \text{ vs. } 1: \quad 119.75 - 57.25 = 62.50 > T_{0.05} = 45.47 \end{aligned}$$

Este análisis indica que con el nivel de temperatura de 70°F, la vida media de la batería es la misma para los materiales tipos 2 y 3, y que la vida media de la batería para el material tipo 1 es significativamente menor.

Si la interacción es significativa, el experimentador podría comparar las medias de *todas* las *ab* celdas para determinar cuáles difieren significativamente. En este análisis, las diferencias entre las medias de las celdas incluyen los efectos de la interacción, así como ambos efectos principales. En el ejemplo 5-1, esto daría 36 comparaciones entre todos los pares posibles de las nueve medias de las celdas.

Salida de computadora

En la figura 5-10 se presenta la salida de computadora de *Design-Expert* para los datos de la vida de la batería del ejemplo 5-1. Observe que

$$\begin{aligned} SS_{\text{Modelo}} &= SS_{\text{Material}} + SS_{\text{Temperatura}} + SS_{\text{Interacción}} \\ &= 10{,}683.72 + 39{,}118.72 + 9613.78 \\ &= 59{,}416.22 \end{aligned}$$

y que

$$R^2 = \frac{SS_{\text{Modelo}}}{SS_{\text{Total}}} = \frac{59{,}416.22}{77{,}646.97} = 0.7652$$

Es decir, cerca de 77% de la variabilidad de la vida de la batería es explicada por el material de la placa de la batería, la temperatura y la interacción entre el tipo de material y la temperatura. En la salida de computadora se muestran también los residuales del modelo ajustado. A continuación se indica cómo usar estos residuales para verificar la adecuación del modelo.

Response: Life in hours
ANOVA for Selected Factorial Model
Analysis of variance table [Partial sum of squares]

Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	59416.22	8	7427.03	11.00	<0.0001	significant
A	*10683.72*	*2*	*5341.86*	*7.91*	*0.0020*	
B	*39118.72*	*2*	*19559.36*	*28.97*	*<0.0001*	
AB	*9613.78*	*4*	*2403.44*	*3.56*	*0.0186*	
Residual	18230.75	27	675.21			
Lack of Fit	*0.000*	*0*				
Pure Error	*18230.75*	*27*	*675.21*			
Cor Total	77646.97	35				

Std. Dev.	25.98	R-Squared	0.7652
Mean	105.53	Adj R-Squared	0.6956
C.V.	24.62	Pred R-Squared	0.5826
PRESS	32410.22	Adeq Precision	8.178

Diagnostics Case Statistics

Standard Order	Actual Value	Predicted Value	Residual	Leverage	Student Residual	Cook's Distance	Outlier t
1	130.00	134.75	−4.75	0.250	−0.211	0.002	−0.207
2	74.00	134.75	−60.75	0.250	−2.700	0.270	−3.100
3	155.00	134.75	20.25	0.250	0.900	0.030	0.897
4	180.00	134.75	45.25	0.250	2.011	0.150	2.140
5	150.00	155.75	−5.75	0.250	−0.256	0.002	−0.251
6	159.00	155.75	3.25	0.250	0.144	0.001	0.142
7	188.00	155.75	32.25	0.250	1.433	0.076	1.463
8	126.00	155.75	−29.75	0.250	−1.322	0.065	−1.341
9	138.00	144.00	−6.00	0.250	−0.267	0.003	−0.262
10	168.00	144.00	24.00	0.250	1.066	0.042	1.069
11	110.00	144.00	−34.00	0.250	−1.511	0.085	−1.550
12	160.00	144.00	16.00	0.250	0.711	0.019	0.704
13	34.00	57.25	−23.25	0.250	−1.033	0.040	−1.035
14	80.00	57.25	22.75	0.250	1.011	0.038	1.011
15	40.00	57.25	−17.25	0.250	−0.767	0.022	−0.761
16	75.00	57.25	17.75	0.250	0.789	0.023	0.783
17	136.00	119.75	16.25	0.250	0.722	0.019	0.716
18	106.00	119.75	−13.75	0.250	−0.611	0.014	−0.604
19	122.00	119.75	2.25	0.250	0.100	0.000	0.098
20	115.00	119.75	−4.75	0.250	−0.211	0.002	−0.207
21	174.00	145.75	28.25	0.250	1.255	0.058	1.269
22	150.00	145.75	4.25	0.250	0.189	0.001	0.185
23	120.00	145.75	−25.75	0.250	−1.144	0.048	−1.151
24	139.00	145.75	−6.75	0.250	−0.300	0.003	−0.295
25	20.00	57.50	−37.50	0.250	−1.666	0.103	−1.726
26	82.00	57.50	24.50	0.250	1.089	0.044	1.093
27	70.00	57.50	12.50	0.250	0.555	0.011	0.548
28	58.00	57.50	0.50	0.250	0.022	0.000	0.022
29	25.00	49.50	−24.50	0.250	−1.089	0.044	−1.093
30	58.00	49.50	8.50	0.250	0.378	0.005	0.372
31	70.00	49.50	20.50	0.250	0.911	0.031	0.908
32	45.00	49.50	−4.50	0.250	−0.200	0.001	−0.196
33	96.00	85.50	10.50	0.250	0.467	0.008	0.460
34	82.00	85.50	−3.50	0.250	−0.156	0.001	−0.153
35	104.00	85.50	18.50	0.250	0.822	0.025	0.817
36	60.00	85.50	−25.50	0.250	−1.133	0.048	−1.139

Figura 5-10 Salida de *Design-Expert* para el ejemplo 5-1.

5-3.3 Verificación de la adecuación del modelo

Antes de adoptar las conclusiones del análisis de varianza, deberá verificarse la adecuación del modelo fundamental. Como anteriormente, la herramienta primaria de diagnóstico es el **análisis residual.** Los residuales del modelo factorial de dos factores son

$$e_{ijk} = y_{ijk} - \hat{y}_{ijk} \tag{5-11}$$

y puesto que el valor ajustado $\hat{y}_{ijk} = \bar{y}_{ij.}$ (el promedio de las observaciones de la celda *ij*-ésima), la ecuación 5-11 queda como

$$e_{ijk} = y_{ijk} - \bar{y}_{ij.} \tag{5-12}$$

En la salida de computadora de *Design-Expert* (figura 5-10) y en la tabla 5-6 se muestran los residuales de los datos de la vida de la batería del ejemplo 5-1. La gráfica de probabilidad normal de estos residuales (figura 5-11) no revela nada particularmente problemático, aun cuando el residual negativo más grande (–60.75 con 15°F para el material tipo 1) se aparta un poco de los demás. El valor estandarizado de este residual es $-60.75/\sqrt{675.21} = -2.34$, y es el único residual cuyo valor absoluto es mayor que 2.

En la figura 5-12 se grafican los residuales contra los valores ajustados $\hat{y}_{ijk}$. Esta gráfica indica una ligera tendencia de la varianza de los residuales a incrementarse cuando la vida de la batería se incrementa. En las figuras 5-13 y 5-14 se grafican los residuales contra los tipos del material y la temperatura, respectivamente. Ambas gráficas indican una ligera desigualdad de la varianza, con la combinación del tratamiento 15°F y material tipo 1, teniendo posiblemente una varianza mayor que las demás.

En la tabla 5-6 se observa que la celda 15°F-material tipo 1 contiene los dos residuales extremos (–60.75 y 45.25). Estos dos residuales son los principales responsables de la desigualdad de la varianza detectada en las figuras 5-12 a 5-14. Al examinarse nuevamente los datos no se observa ningún problema obvio, tal como un error al registrar los datos, por lo que estas respuestas se aceptan como legítimas. Es posible que esta combinación de tratamientos particular produzca una vida de la batería ligeramente más errática que las demás. Sin embargo, el problema no es lo suficientemente grave como para tener un impacto dramático en el análisis y las conclusiones.

5-3.4 Estimación de los parámetros del modelo

Los parámetros del modelo de los efectos para el diseño factorial de dos factores

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \varepsilon_{ijk} \tag{5-13}$$

Tabla 5-6 Residuales del ejemplo 5-1

Tipo de material	Temperatura (°F)					
	15		70		125	
1	–4.75	20.25	–23.25	–17.25	–37.50	12.50
	–60.75	45.25	22.75	17.75	24.50	0.50
2	–5.75	32.25	16.25	2.25	–24.50	20.50
	3.25	–29.75	–13.75	–4.75	8.50	–4.50
3	–6.00	–34.00	28.25	–25.75	10.50	18.50
	24.00	16.00	4.25	–6.75	–3.50	–25.50

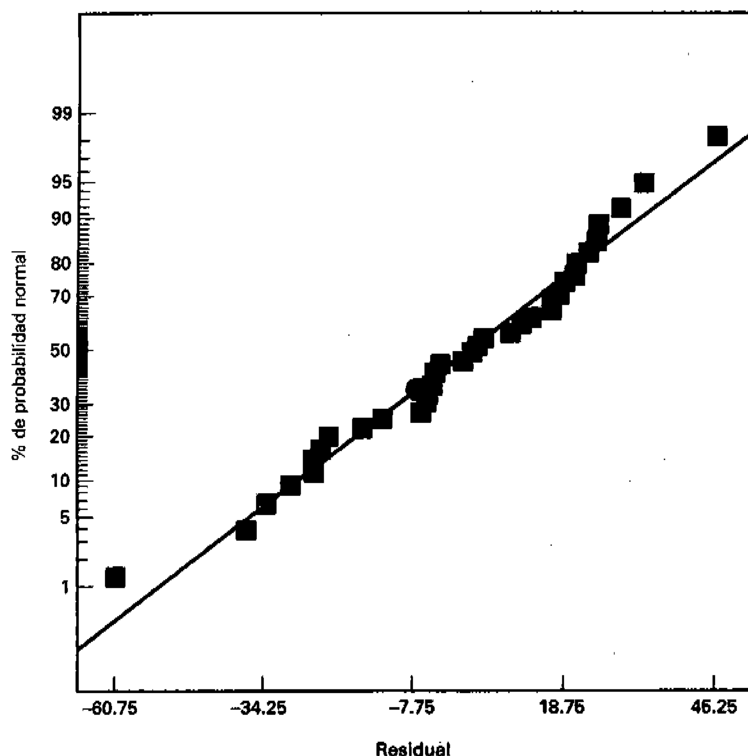


Figura 5-11 Gráfica de probabilidad normal de los residuales del ejemplo 5-1.

pueden estimarse por mínimos cuadrados. Puesto que el modelo tiene $1 + a + ab$ parámetros que deben estimarse, hay $1 + a + b + ab$ ecuaciones normales. Al utilizar el método de la sección 3-9, no es difícil demostrar que las ecuaciones normales son

$$\mu\colon abn\hat{\mu} + bn\sum_{i=1}^{a} \hat{\tau}_i + an\sum_{j=1}^{b}\hat{\beta}_j + n\sum_{i=1}^{a}\sum_{j=1}^{b} (\widehat{\tau\beta})_{ij} = y_{...} \tag{5-14a}$$

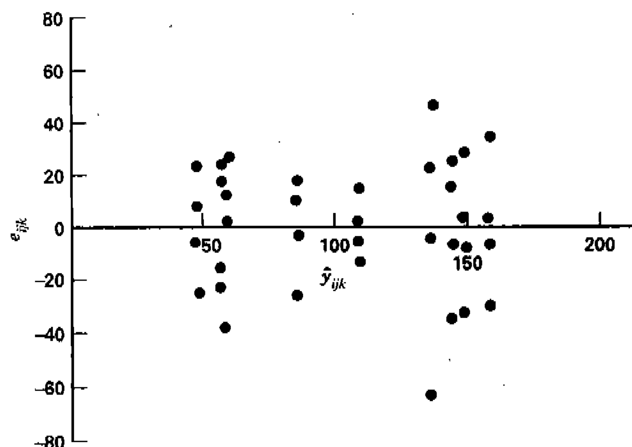


Figura 5-12 Gráfica de los residuales contra $\hat{y}_{ijk}$ para el ejemplo 5-1.

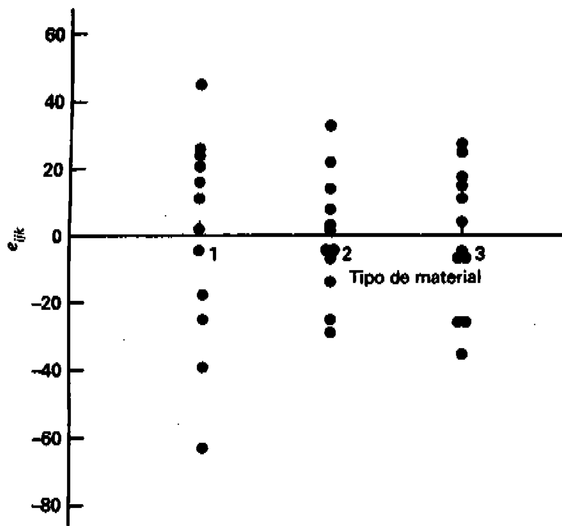


Figura 5-13 Gráfica de los residuales contra el tipo de material para el ejemplo 5-1.

$$\tau_i : bn\hat{\mu} + bn\hat{\tau}_i + n\sum_{j=1}^{b} \hat{\beta}_j + n\sum_{j=1}^{b} (\widehat{\tau\beta})_{ij} = y_{i..} \qquad i = 1, 2, \ldots, a \tag{5-14b}$$

$$\beta_j : an\hat{\mu} + n\sum_{i=1}^{a} \hat{\tau}_i + an\hat{\beta}_j + n\sum_{i=1}^{a} (\widehat{\tau\beta})_{ij} = y_{.j.} \qquad j = 1, 2, \ldots, b \tag{5-14c}$$

$$(\tau\beta)_{ij} : n\hat{\mu} + n\hat{\tau}_i + n\hat{\beta}_j + n(\widehat{\tau\beta})_{ij} = y_{ij.} \qquad \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, b \end{cases} \tag{5-14d}$$

Por conveniencia, el parámetro que corresponde a cada ecuación normal se indica a la izquierda de las ecuaciones 5-14.

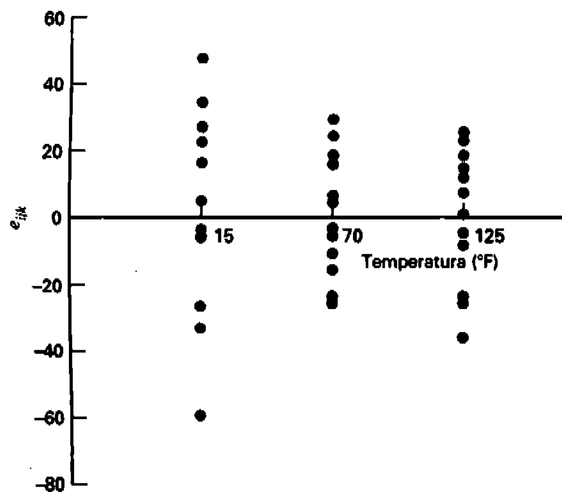


Figura 5-14 Gráfica de los residuales contra la temperatura para el ejemplo 5-1.

El modelo de los efectos (ecuación 5-13) está sobreparametrizado. Observe que la suma de las a ecuaciones de la ecuación 5-14b es igual a la ecuación 5-14a y que la suma de las b ecuaciones de la ecuación 5-14c es igual a la ecuación 5-14a. Asimismo, la operación suma de la ecuación 5-14d sobre j para una i particular dará la ecuación 5-14b, y la operación suma de la ecuación 5-14d sobre i para una j particular dará la ecuación 5-14c. Por lo tanto, hay $a + b + 1$ **dependencias lineales** en este sistema de ecuaciones y no existirá ninguna solución única. A fin de obtener una solución, se imponen las restricciones

$$\sum_{i=1}^{a} \hat{\tau}_i = 0 \tag{5-15a}$$

$$\sum_{j=1}^{b} \hat{\beta}_j = 0 \tag{5-15b}$$

$$\sum_{i=1}^{a} (\widehat{\tau\beta})_{ij} = 0 \qquad j = 1, 2, \ldots, b \tag{5-15c}$$

y

$$\sum_{j=1}^{b} (\widehat{\tau\beta})_{ij} = 0 \qquad i = 1, 2, \ldots, a \tag{5-15d}$$

Las ecuaciones 5-15a y 5-15b constituyen dos restricciones, mientras que las ecuaciones 5-15c y 5-15d forman $a + b - 1$ restricciones independientes. Por lo tanto, se tienen en total $a + b + 1$ restricciones, el número que se requiere.

Al aplicar estas restricciones, las ecuaciones normales (ecuaciones 5-14) se simplifican considerablemente, y se obtiene la solución

$$\begin{aligned}
\hat{\mu} &= \bar{y}_{...} \\
\hat{\tau}_i &= \bar{y}_{i..} - \bar{y}_{...} \qquad i = 1, 2, \ldots, a \\
\hat{\beta}_j &= \bar{y}_{.j.} - \bar{y}_{...} \qquad j = 1, 2, \ldots, b \\
(\widehat{\tau\beta})_{ij} &= \bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...} \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, b \end{cases}
\end{aligned} \tag{5-16}$$

Observe el gran atractivo intuitivo de esta solución de las ecuaciones normales. Los efectos de los tratamientos de los renglones se estiman con el promedio del renglón menos el gran promedio; los tratamientos de las columnas se estiman con el promedio de la columna menos el gran promedio, y la interacción ij-ésima se estima con el promedio de la celda ij-ésima menos el gran promedio, el efecto del renglón i-ésimo y el efecto de la columna j-ésima.

Al utilizar la ecuación 5-16, el **valor ajustado** y_{ijk} puede encontrarse como

$$\begin{aligned}
\hat{y}_{ijk} &= \hat{\mu} + \hat{\tau}_i + \hat{\beta}_j + (\widehat{\tau\beta})_{ij} \\
&= \bar{y}_{...} + (\bar{y}_{i..} - \bar{y}_{...}) + (\bar{y}_{.j.} - \bar{y}_{...}) \\
&\quad + (\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...}) \\
&= \bar{y}_{ij.}
\end{aligned}$$

Es decir, la observación k-ésima de la celda ij-ésima se estima con el promedio de las n observaciones de esa celda. Este resultado se usó en la ecuación 5-12 para obtener los residuales del modelo factorial de dos factores.

Puesto que se han usado restricciones (ecuaciones 5-15) para resolver las ecuaciones normales, los parámetros del modelo no tienen estimaciones únicas. Sin embargo, ciertas **funciones** importantes de los parámetros del modelo *son* estimables, es decir, tienen una estimación única independientemente de las restricciones elegidas. Un ejemplo es $\tau_i - \tau_u + (\overline{\tau\beta})_{i.} - (\overline{\tau\beta})_{u.}$, que podría considerarse como la "verdade-

ra" diferencia entre los niveles i-ésimo y u-ésimo del factor A. Observe que la verdadera diferencia entre los niveles de cualquier efecto principal incluye un efecto de la interacción "promedio". Es este resultado el que perturba las pruebas de los efectos principales en presencia de una interacción, como se señaló anteriormente. En general, cualquier función de los parámetros del modelo que sea una combinación lineal del miembro izquierdo de las ecuaciones normales es **estimable**. Esta propiedad también se hizo notar en el capítulo 3 cuando se estudió el modelo de un solo factor. Para mayores detalles, ver el material suplementario del texto de este capítulo.

5-3.5 Elección del tamaño de la muestra

Para determinar un tamaño de la muestra (el número de réplicas, n) apropiado en un diseño factorial de dos factores, el experimentador puede apoyarse en las curvas de operación característica que aparecen en la parte V del apéndice. En la tabla 5-7 se muestra el valor apropiado del parámetro Φ^2, así como los grados de libertad del numerador y el denominador.

Una forma muy eficaz de emplear estas curvas consiste en encontrar el valor menor de Φ^2 que corresponde a una diferencia especificada entre las medias de dos tratamientos cualesquiera. Por ejemplo, si la diferencia en las medias de dos renglones cualesquiera es D, entonces el valor mínimo de Φ^2 es

$$\Phi^2 = \frac{nbD^2}{2a\sigma^2} \tag{5-17}$$

mientras que si la diferencia en las medias de dos columnas cualesquiera es D, entonces el valor mínimo de Φ^2 es

$$\Phi^2 = \frac{naD^2}{2b\sigma^2} \tag{5-18}$$

Por último, el valor mínimo de Φ^2 que corresponde a una diferencia D entre dos efectos de interacción cualesquiera es

$$\Phi^2 = \frac{nD^2}{2\sigma^2[(a-1)(b-1)+1]} \tag{5-19}$$

Para ilustrar el uso de estas ecuaciones, considere los datos de la vida de la batería del ejemplo 5-1. Suponga que antes de correr el experimento se decide que la hipótesis nula deberá rechazarse con una alta

Tabla 5-7 Parámetros de la curva de operación característica de la parte V del apéndice para el diseño factorial de dos factores, modelo con efectos fijos

Factor	Φ^2	Grados de libertad del numerador	Grados de libertad del denominador
A	$\dfrac{bn\sum_{i=1}^{a}\tau_i^2}{a\sigma^2}$	$a-1$	$ab(n-1)$
B	$\dfrac{an\sum_{j=1}^{b}\beta_j^2}{b\sigma^2}$	$b-1$	$ab(n-1)$
AB	$\dfrac{n\sum_{i=1}^{a}\sum_{j=1}^{b}(\tau\beta)_{ij}^2}{\sigma^2[(a-1)(b-1)+1]}$	$(a-1)(b-1)$	$ab(n-1)$

probabilidad si la diferencia en la vida media de la batería entre dos temperaturas cualesquiera es hasta de 40 horas. Por lo tanto $D = 40$, y si se supone que la desviación estándar de la vida de la batería es aproximadamente 25, entonces por la ecuación 5-18 se obtiene

$$\Phi^2 = \frac{naD^2}{2b\sigma^2} = \frac{n(3)(40)^2}{2(3)(25)^2} = 1.28n$$

como el valor mínimo de Φ^2. Suponiendo que $\alpha = 0.05$, ahora puede usarse la parte V del apéndice para construir la tabla siguiente:

n	Φ^2	Φ	ν_1 = Grados de libertad del numerador	ν_2 = Grados de libertad del error	β
2	2.56	1.60	2	9	0.45
3	3.84	1.96	2	18	0.18
4	5.12	2.26	2	27	0.06

Observe que con $n = 4$ réplicas se obtiene un riesgo β de cerca de 0.06, o una probabilidad aproximada de 94%, de rechazar la hipótesis nula si la diferencia en la vida media de la batería con dos niveles de temperatura cualesquiera es hasta de 40 horas. Por lo tanto, se concluye que cuatro réplicas bastan para proporcionar la sensitividad deseada siempre y cuando la estimación usada para la desviación estándar de la vida de la batería no tenga un error grave. En caso de duda, el experimentador podría repetir el procedimiento anterior con otros valores de σ para determinar el efecto que tendría una estimación equivocada de este parámetro sobre la sensitividad del diseño.

5-3.6 El supuesto de no interacción en un modelo de dos factores

Ocasionalmente, un experimentador siente que es apropiado un **modelo de dos factores sin interacción,** por ejemplo

$$y_{ijk} = \mu + \tau_i + \beta_j + \varepsilon_{ijk} \qquad \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, b \\ k = 1, 2, \ldots, n \end{cases} \tag{5-20}$$

Sin embargo, se deberá ser muy cuidadoso al hacer caso omiso de los términos de interacción, ya que la presencia de una interacción significativa puede tener un impacto dramático sobre la interpretación de los datos.

El análisis estadístico de un modelo factorial de dos factores sin interacción es directo. En la tabla 5-8 se presenta el análisis de los datos de la vida de la batería del ejemplo 5-1, suponiendo que es válido el mo-

Tabla 5-8 Análisis de varianza de los datos de la vida de la batería suponiendo que no hay interacción

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0
Tipos de material	10,683.72	2	5,341.86	5.95
Temperatura	39,118.72	2	19,559.36	21.78
Error	27,844.52	31	898.21	
Total	77,646.96	35		

delo sin interacción (ecuación 5-20). Como ya se señaló, los dos efectos principales son significativos. Sin embargo, tan pronto como se efectúa el análisis residual de estos datos, se pone de manifiesto que el modelo sin interacción es inadecuado. Para el modelo de dos factores sin interacción, los valores ajustados son $\hat{y}_{ijk} = \bar{y}_{i..} + \bar{y}_{.j.} - \bar{y}_{...}$. En la figura 5-15 se presenta la gráfica de $\bar{y}_{ij.} - \hat{y}_{ijk}$ (los promedios de las celdas menos el valor ajustado de esa celda) contra el valor ajustado $\hat{y}_{ijk}$. Ahora las cantidades $\bar{y}_{ij.} - \hat{y}_{ijk}$ pueden considerarse como las diferencias entre las medias de las celdas observadas y las medias de las celdas estimadas suponiendo que no hay interacción. Cualquier patrón en estas cantidades sugiere la presencia de una interacción. En la figura 5-15 se observa un patrón claro cuando las cantidades $\bar{y}_{ij.} - \hat{y}_{ijk}$ pasan de positivo a negativo, y después de nuevo a positivo y a negativo. Esta estructura es el resultado de la interacción entre los tipos del material y la temperatura.

5-3.7 Una observación por celda

En ocasiones se encuentran experimentos de dos factores con una **sola réplica,** es decir, en los que sólo hay una observación por celda. Cuando hay dos factores y una sola observación por celda, el modelo de los efectos es

$$y_{ij} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \varepsilon_{ij} \qquad \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, b \end{cases} \tag{5-21}$$

El análisis de varianza para esta situación se presenta en la tabla 5-9, suponiendo que ambos factores son fijos.

Al examinar los cuadrados medios esperados, se observa que la varianza del error σ^2 es *no estimable;* es decir, que el efecto de la interacción de los dos factores $(\tau\beta)_{ij}$ y el error experimental no pueden separarse de alguna manera obvia. Por consiguiente, no se cuenta con pruebas para los efectos principales a menos que el efecto de la interacción sea cero. Si no hay una interacción presente, entonces $(\tau\beta)_{ij} = 0$ para toda i y j, y un modelo plausible es

$$y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij} \qquad \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, b \end{cases} \tag{5-22}$$

Si el modelo (ecuación 5-22) es apropiado, entonces el cuadrado medio de los residuales de la tabla 5-9 es un estimador insesgado de σ^2, y los efectos principales pueden probarse comparando MS_A y MS_B con MS_{Residual}.

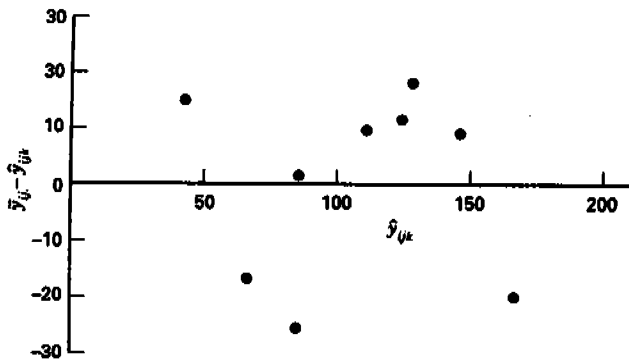


Figura 5-15 Gráfica de $\bar{y}_{ij.} - \hat{y}_{ijk}$ contra $\hat{y}_{ijk}$ para los datos de la vida de la batería.

Tabla 5-9 Análisis de varianza de un modelo de dos factores, una observación por celda

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	Cuadrado medio esperado
Renglones (A)	$\sum_{i=1}^{a} \frac{y_{i.}^2}{b} - \frac{y_{..}^2}{ab}$	$a-1$	MS_A	$\sigma^2 + \frac{b\sum \tau_i^2}{a-1}$
Columnas (B)	$\sum_{j=1}^{b} \frac{y_{.j}^2}{a} - \frac{y_{..}^2}{ab}$	$b-1$	MS_B	$\sigma^2 + \frac{a\sum \beta_j^2}{b-1}$
Residual o AB	Sustracción	$(a-1)(b-1)$	MS_{Residual}	$\sigma^2 + \frac{\sum\sum(\tau\beta)_{ij}^2}{(a-1)(b-1)}$
Total	$\sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij}^2 - \frac{y_{..}^2}{ab}$	$ab-1$		

Una prueba desarrollada por Tukey [111a] es útil para determinar si está presente una interacción. En el procedimiento se supone que el término de la interacción tiene una forma particularmente simple, a saber,

$$(\tau\beta)_{ij} = \gamma\tau_i\beta_j$$

donde γ es una constante desconocida. Al definir así el término de la interacción, puede usarse un enfoque de regresión para probar la significación del término de la interacción. En la prueba se hace la partición de la suma de cuadrados de los residuales en un componente con un solo grado de libertad debido a la no aditividad (interacción) y un componente del error con $(a-1)(b-1)-1$ grados de libertad. En lo que a los cálculos se refiere, se tiene

$$SS_N = \frac{\left[\sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij}y_{i.}y_{.j} - y_{..}\left(SS_A + SS_B + \frac{y_{..}^2}{ab}\right)\right]^2}{abSS_ASS_B} \tag{5-23}$$

con un grado de libertad, y

$$SS_{\text{Error}} = SS_{\text{Residual}} - SS_N \tag{5-24}$$

con $(a-1)(b-1)-1$ grados de libertad. Para probar la presencia de una interacción, se calcula

$$F_0 = \frac{SS_N}{SS_{\text{Error}}/[(a-1)(b-1)-1]} \tag{5-25}$$

Si $F_0 > F_{\alpha, 1, (a-1)(b-1)-1}$, debe rechazarse la hipótesis de que no hay ninguna interacción.

EJEMPLO 5-2 ....

Las impurezas presentes en un producto químico son afectadas por dos factores, la presión y la temperatura. En la tabla 5-10 se muestran los datos de una sola réplica de un experimento factorial. Las sumas de cuadrados son

$$SS_A = \frac{1}{b}\sum_{i=1}^{a} y_{i.}^2 - \frac{y_{..}^2}{ab}$$
$$= \frac{1}{5}[23^2 + 13^2 + 8^2] - \frac{44^2}{(3)(5)} = 23.33$$

Tabla 5-10 Datos de las impurezas del ejemplo 5-2

Temperatura (°F)	Presión					
	25	30	35	40	45	$y_{i.}$
100	5	4	6	3	5	23
125	3	1	4	2	3	13
150	1	1	3	1	2	8
$y_{.j}$	9	6	13	6	10	$44 = y_{..}$

$$SS_B = \frac{1}{a}\sum_{j=1}^{b} y_{.j}^2 - \frac{y_{..}^2}{ab}$$

$$= \frac{1}{3}[9^2 + 6^2 + 13^2 + 6^2 + 10^2] - \frac{44^2}{(3)(5)} = 11.60$$

$$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij}^2 - \frac{y_{..}^2}{ab}$$

$$= 166 - 129.07 = 36.93$$

y

$$SS_{\text{Residual}} = SS_T - SS_A - SS_B$$

$$= 36.96 - 23.33 - 11.60 = 2.00$$

La suma de cuadrados de la no aditividad se calcula con la ecuación 5-23 de la siguiente manera:

$$\sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij}y_{i.}y_{.j} = (5)(23)(9) + (4)(23)(6) + \cdots + (2)(8)(10) = 7236$$

$$SS_N = \frac{\left[\sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij}y_{i.}y_{.j} - y_{..}\left(SS_A + SS_B + \frac{y_{..}^2}{ab}\right)\right]^2}{abSS_A SS_B}$$

$$= \frac{[7236 - (44)(23.33 + 11.60 + 129.07)]^2}{(3)(5)(23.33)(11.60)}$$

$$= \frac{[20.00]^2}{4059.42} = 0.0985$$

y la suma de cuadrados del error es, por la ecuación 5-24,

$$SS_{\text{Error}} = SS_{\text{Residual}} - SS_N$$

$$= 2.00 - 0.0985 = 1.9015$$

El análisis de varianza completo se resume en la tabla 5-11. El estadístico de prueba para la no aditividad es $F_0 = 0.0985/0.2716 = 0.36$, de donde se concluye que no hay evidencia de interacción en estos datos. Los efectos principales de la temperatura y la presión son significativos.

Para concluir esta sección, se hace notar que el modelo factorial de dos factores con una observación por celda (ecuación 5-22) luce exactamente igual que el modelo de bloques completos aleatorizados

Tabla 5-11 Análisis de varianza del ejemplo 5-2

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0	Valor *P*
Temperatura	23.33	2	11.67	42.97	0.0001
Presión	11.60	4	2.90	10.68	0.0042
No aditividad	0.0985	1	0.0985	0.36	0.5674
Error	1.9015	7	0.2716		
Total	36.93	14			

(ecuación 4-1). De hecho, la prueba de Tukey con un solo grado de libertad para la no aditividad puede aplicarse directamente para probar la presencia de una interacción en el modelo de bloques aleatorizados. Sin embargo, es necesario recordar que las **situaciones experimentales** que llevan al modelo de bloques aleatorizados y al modelo factorial son muy diferentes. En el modelo factorial, *todas* las *ab* corridas se hacen de manera aleatoria, mientras que en el modelo de bloques aleatorizados la aleatorización sólo ocurre *dentro del bloque*. Los bloques constituyen una restricción sobre la aleatorización. Por lo tanto, la manera en que se corren los experimentos, así como la interpretación de los dos modelos, es muy diferente.

5-4 DISEÑO FACTORIAL GENERAL

Los resultados del diseño factorial de dos factores pueden ampliarse al caso general en que hay *a* niveles del factor *A*, *b* niveles del factor *B*, *c* niveles del factor *C*, etc., dispuestos en un experimento factorial. En general, habrá $abc \cdots n$ observaciones totales si se hacen *n* réplicas del experimento completo. De nueva cuenta, observe que es necesario un mínimo de dos réplicas ($n \geq 2$) para determinar una suma de cuadrados debida al error si todas las interacciones posibles están incluidas en el modelo.

Cuando todos los factores del experimento son fijos, es sencillo formular y probar hipótesis acerca de los efectos principales y las interacciones. Para un modelo con efectos fijos, los estadísticos de prueba para cada efecto principal e interacción pueden construirse dividiendo el cuadrado medio correspondiente del efecto o interacción por el cuadrado medio del error. Todas estas pruebas *F* serán de una cola superior. El número de grados de libertad de cualquier efecto principal es el número de niveles del factor menos uno, y el número de grados de libertad de una interacción es el producto del número de grados de libertad asociados con los componentes individuales de la interacción.

Por ejemplo, considere el **modelo del análisis de varianza de tres factores:**

$$y_{ijkl} = \mu + \tau_i + \beta_j + \gamma_k + (\tau\beta)_{ij} + (\tau\gamma)_{ik} + (\beta\gamma)_{jk} + (\tau\beta\gamma)_{ijk} + \varepsilon_{ijkl} \quad \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, b \\ k = 1, 2, \ldots, c \\ l = 1, 2, \ldots, n \end{cases} \tag{5-26}$$

Suponiendo que *A*, *B* y *C* son fijos, la **tabla del análisis de varianza** se presenta en la tabla 5-12. Las pruebas *F* para los efectos principales y las interacciones se siguen directamente de los cuadrados medios esperados.

Tabla 5-12 La tabla del análisis de varianza del modelo de tres factores con efectos fijos

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	Cuadrado medio esperado	F_0
A	SS_A	$a-1$	MS_A	$\sigma+\frac{bcn\sum \tau_i^2}{a-1}$	$F_0=\frac{MS_A}{MS_E}$
B	SS_B	$b-1$	MS_B	$\sigma^2+\frac{acn\sum \beta_j^2}{b-1}$	$F_0=\frac{MS_B}{MS_E}$
C	SS_C	$c-1$	MS_C	$\sigma^2+\frac{abn\sum \gamma_k^2}{c-1}$	$F_0=\frac{MS_C}{MS_E}$
AB	SS_{AB}	$(a-1)(b-1)$	MS_{AB}	$\sigma^2+\frac{cn\sum\sum (\tau\beta)_{ij}^2}{(a-1)(b-1)}$	$F_0=\frac{MS_{AB}}{MS_E}$
AC	SS_{AC}	$(a-1)(c-1)$	MS_{AC}	$\sigma^2+\frac{bn\sum\sum (\tau\gamma)_{ik}^2}{(a-1)(c-1)}$	$F_0=\frac{MS_{AC}}{MS_E}$
BC	SS_{BC}	$(b-1)(c-1)$	MS_{BC}	$\sigma^2+\frac{an\sum\sum (\beta\gamma)_{jk}^2}{(b-1)(c-1)}$	$F_0=\frac{MS_{BC}}{MS_E}$
ABC	SS_{ABC}	$(a-1)(b-1)(c-1)$	MS_{ABC}	$\sigma^2+\frac{n\sum\sum\sum (\tau\beta\gamma)_{ijk}^2}{(a-1)(b-1)(c-1)}$	$F_0=\frac{MS_{ABC}}{MS_E}$
Error	SS_E	$abc(n-1)$	MS_E	σ^2	
Total	SS_T	$abcn-1$			

En general, los cálculos del análisis de varianza se efectuarían utilizando un paquete de software de estadística. Sin embargo, en ocasiones resultan útiles las fórmulas para calcular manualmente las sumas de cuadrados de la tabla 5-12. La suma de cuadrados total se encuentra de la manera acostumbrada como

$$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}\sum_{l=1}^{n} y_{ijkl}^2 - \frac{y_{....}^2}{abcn} \tag{5-27}$$

Las sumas de cuadrados de los efectos principales se encuentran a partir de los totales de los factores $A(y_{i...})$, $B(y_{.j..})$ y $C(y_{..k.})$ de la siguiente manera:

$$SS_A = \frac{1}{bcn}\sum_{i=1}^{a} y_{i...}^2 - \frac{y_{....}^2}{abcn} \tag{5-28}$$

$$SS_B = \frac{1}{acn}\sum_{j=1}^{b} y_{.j..}^2 - \frac{y_{....}^2}{abcn} \tag{5-29}$$

$$SS_C = \frac{1}{abn}\sum_{k=1}^{c} y_{..k.}^2 - \frac{y_{....}^2}{abcn} \tag{5-30}$$

Para calcular las sumas de cuadrados de las interacciones de dos factores, se necesitan los totales de las celdas $A \times B, A \times C$ y $B \times C$. Con frecuencia es útil desplegar la tabla de los datos originales en tres tablas de dos vías para calcular estas cantidades. Las sumas de cuadrados se encuentran con

$$\begin{aligned} SS_{AB} &= \frac{1}{cn}\sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij..}^2 - \frac{y_{....}^2}{abcn} - SS_A - SS_B \\ &= SS_{\text{Subtotales}(AB)} - SS_A - SS_B \end{aligned} \tag{5-31}$$

$$\begin{aligned} SS_{AC} &= \frac{1}{bn}\sum_{i=1}^{a}\sum_{k=1}^{c} y_{i.k.}^2 - \frac{y_{....}^2}{abcn} - SS_A - SS_C \\ &= SS_{\text{Subtotales}(AC)} - SS_A - SS_C \end{aligned} \tag{5-32}$$

y

$$\begin{aligned} SS_{BC} &= \frac{1}{an}\sum_{j=1}^{b}\sum_{k=1}^{c} y_{.jk.}^2 - \frac{y_{....}^2}{abcn} - SS_B - SS_C \\ &= SS_{\text{Subtotales}(BC)} - SS_B - SS_C \end{aligned} \tag{5-33}$$

Observe que las sumas de cuadrados de los subtotales de dos factores se encuentran a partir de los totales de cada tabla de dos vías. La suma de cuadrados de la interacción de los tres factores se calcula a partir de los totales de las celdas $\{y_{ijk.}\}$ de tres vías como

$$SS_{ABC} = \frac{1}{n}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c} y_{ijk.}^2 - \frac{y_{....}^2}{abcn} - SS_A - SS_B - SS_C - SS_{AB} - SS_{AC} - SS_{BC} \tag{5-34a}$$

$$= SS_{\text{Subtotales}(ABC)} - SS_A - SS_B - SS_C - SS_{AB} - SS_{AC} - SS_{BC} \tag{5-34b}$$

La suma de cuadrados del error puede encontrarse restando la suma de cuadrados de cada efecto principal e interacción de la suma de cuadrados total o con

$$SS_E = SS_T - SS_{\text{Subtotales}(ABC)} \tag{5-35}$$

EJEMPLO 5-3 ··

El problema del embotellado de un refresco

Una empresa embotelladora de refrescos está interesada en obtener alturas de llenado más uniformes en las botellas que se fabrican en su proceso de manufactura. Teóricamente, la máquina de llenado llena cada botella a la altura objetivo correcta, pero en la práctica, existe variación en torno a este objetivo, y a la embotelladora le gustaría entender mejor las fuentes de esta variabilidad y, en última instancia, reducirla.

El ingeniero del proceso puede controlar tres variables durante el proceso de llenado: el porcentaje de carbonatación (A), la presión de operación en el llenador (B) y las botellas producidas por minuto o rapidez de línea (C). Es sencillo controlar la presión y la rapidez, pero el porcentaje de carbonatación es más difícil de controlar durante la manufactura real debido a que varía con la temperatura. Sin embargo, para los fines de un experimento, el ingeniero puede controlar la carbonatación en tres niveles: 10, 12 y 14 por ciento. Elige dos niveles para la presión (25 y 30 psi) y dos niveles para la rapidez de línea (200 y 250 bpm). El ingeniero decide correr dos réplicas de un diseño factorial con estos tres factores, haciendo las 24 corridas de manera aleatoria. La variable de respuesta observada es la desviación promedio de la altura del llenado objetivo que se observa en una corrida de producción de botellas con cada conjunto de condiciones. En la tabla 5-13 se muestran los datos que resultaron de este experimento. Las desviaciones positivas son alturas de llenado arriba del objetivo, mientras que las desviaciones negativas son alturas de llenado abajo del objetivo. Los números encerrados en círculos de la tabla 5-13 son los totales de las celdas de tres vías $y_{ijk.}$.

La suma de cuadrados total corregida que se encuentra con la ecuación 5-27 es

$$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}\sum_{l=1}^{n} y_{ijkl}^2 - \frac{y_{....}^2}{abcn}$$

$$= 571 - \frac{(75)^2}{24} = 336.625$$

Tabla 5-13 Datos de la desviación de la altura de llenado del ejemplo 5-3

Porcentaje de carbonatación (A)	Presión de operación (B): 25 psi, Rapidez de línea (C): 200		25 psi, 250		30 psi, Rapidez de línea (C): 200		30 psi, 250		$y_{i...}$
10	−3, −1	(−4)	−1, 0	(−1)	−1, 0	(−1)	1, 1	(2)	−4
12	0, 1	(1)	2, 1	(3)	2, 3	(5)	6, 5	(11)	20
14	5, 4	(9)	7, 6	(13)	7, 9	(16)	10, 11	(21)	59
Totales $B \times C$ $y_{.jk.}$	6		15		20		34		$75 = y_{....}$
$y_{.j..}$	21				54				

Totales $A \times B$

$y_{ij..}$

A	B: 25	30
10	−5	1
12	4	16
14	22	37

Totales $A \times C$

$y_{i.k.}$

A	C: 200	250
10	−5	1
12	6	14
14	25	34

y las sumas de cuadrados de los efectos principales que se calculan con las ecuaciones 5-28, 5-29 y 5-30 son

$$SS_{\text{Carbonatación}} = \frac{1}{bcn}\sum_{i=1}^{a} y_{i...}^2 - \frac{y_{....}^2}{abcn}$$

$$= \frac{1}{8}[(-4)^2 + (20)^2 + (59)^2] - \frac{(75)^2}{24} = 252.750$$

$$SS_{\text{Presión}} = \frac{1}{acn}\sum_{j=1}^{b} y_{.j..}^2 - \frac{y_{....}^2}{abcn}$$

$$= \frac{1}{12}[(21)^2 + (54)^2] - \frac{(75)^2}{24} = 45.375$$

y

$$SS_{\text{Rapidez}} = \frac{1}{abn}\sum_{k=1}^{c} y_{..k.}^2 - \frac{y_{....}^2}{abcn}$$

$$= \frac{1}{12}[(26)^2 + (49)^2] - \frac{(75)^2}{24} = 22.042$$

Para calcular las sumas de cuadrados de las interacciones de dos factores, se deben encontrar los totales de las celdas de dos vías. Por ejemplo, para encontrar la carbonatación-presión o interacción AB, se necesitan los totales de las celdas $A \times B$ $\{y_{ij..}\}$ que se muestran en la tabla 5-13. Utilizando la ecuación 5-31, se encuentra que las sumas de cuadrados son

$$SS_{AB} = \frac{1}{cn}\sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij..}^2 - \frac{y_{....}^2}{abcn} - SS_A - SS_B$$

$$= \frac{1}{4}[(-5)^2 + (1)^2 + (4)^2 + (16)^2 + (22)^2 + (37)^2] - \frac{(75)^2}{24} - 252.750 - 45.375$$

$$= 5.250$$

Para la carbonatación-rapidez o interacción AC se usan los totales de las celdas $A \times C$ $\{y_{i.k.}\}$ que se muestran en la tabla 5-13 y la ecuación 5-32:

$$SS_{AC} = \frac{1}{bn}\sum_{i=1}^{a}\sum_{k=1}^{c} y_{i.k.}^2 - \frac{y_{....}^2}{abcn} - SS_A - SS_C$$

$$= \frac{1}{4}[(-5)^2 + (1)^2 + (6)^2 + (14)^2 + (25)^2 + (34)^2] - \frac{(75)^2}{24} - 252.750 - 22.042$$

$$= 0.583$$

La presión-rapidez o interacción BC se encuentra con los totales de las celdas $B \times C$ $\{y_{.jk.}\}$ que se muestran en la tabla 5-13 y la ecuación 5-33:

$$SS_{BC} = \frac{1}{an}\sum_{j=1}^{b}\sum_{k=1}^{c} y_{.jk.}^2 - \frac{y_{....}^2}{abcn} - SS_B - SS_C$$

$$= \frac{1}{6}[(6)^2 + (15)^2 + (20)^2 + (34)^2] - \frac{(75)^2}{24} - 45.375 - 22.042$$

$$= 1.042$$

La suma de cuadrados de la interacción de los tres factores se encuentra con los totales de las celdas $A \times B \times C$ $\{y_{ijk.}\}$, los cuales están encerrados en un círculo en la tabla 5-13. Por la ecuación 5-34a se encuentra

$$
\begin{aligned}
SS_{ABC} &= \frac{1}{n}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c} y_{ijk.}^2 - \frac{y_{....}^2}{abcn} - SS_A - SS_B - SS_C - SS_{AB} - SS_{AC} - SS_{BC} \\
&= \frac{1}{2}[(-4)^2 + (-1)^2 + (-1)^2 + \cdots + (16)^2 + (21)^2] - \frac{(75)^2}{24} \\
&\quad - 252.750 - 45.375 - 22.042 - 5.250 - 0.583 - 1.042 \\
&= 1.083
\end{aligned}
$$

Por último, al observar que

$$
SS_{\text{Subtotales}(ABC)} = \frac{1}{n}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c} y_{ijk.}^2 - \frac{y_{....}^2}{abcn} = 328.125
$$

se tiene

$$
\begin{aligned}
SS_E &= SS_T - SS_{\text{Subtotales}(ABC)} \\
&= 336.625 - 328.125 \\
&= 8.500
\end{aligned}
$$

En la tabla 5-14 se resume el análisis de varianza. Se observa que el porcentaje de carbonatación, la presión de operación y la rapidez de línea afectan significativamente el volumen de llenado. El cociente F de la interacción carbonatación-presión tiene un valor P de 0.0558, lo cual indica cierta interacción entre estos factores.

El siguiente paso deberá ser un análisis de los residuales de este experimento. Se deja como ejercicio para el lector, pero se señala que la gráfica de probabilidad normal de los residuales y los demás diagnósticos usuales no indican ningún motivo de preocupación importante.

Como ayuda para la interpretación práctica de este experimento, en la figura 5-16 se grafican los tres efectos principales y la interacción AB (carbonatación-presión). Las representaciones de los efectos principales son sólo gráficas de los promedios de las respuestas marginales para los niveles de los tres factores. Observe que las tres variables tienen efectos principales *positivos;* es decir, el incremento de la variable mueve hacia arriba la desviación promedio del llenado objetivo. La interacción entre la cabonatación y la presión es bastante pequeña, como lo indica la forma similar de las dos curvas de la figura 5-16*d*.

Puesto que la empresa quiere que la desviación promedio del llenado objetivo esté cerca de cero, el ingeniero decide recomendar el nivel bajo de la presión de operación (25 psi) y el nivel alto de la rapidez de línea (250 bpm, que maximizará la rapidez de producción). En la figura 5-17 se grafica la desviación

Tabla 5-14 Análisis de varianza del ejemplo 5-3

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0	Valor P
Porcentaje de carbonatación (A)	252.750	2	126.375	178.412	<0.0001
Presión de operación (B)	45.375	1	45.375	64.059	<0.0001
Rapidez de línea (C)	22.042	1	22.042	31.118	0.0001
AB	5.250	2	2.625	3.706	0.0558
AC	0.583	2	0.292	0.412	0.6713
BC	1.042	1	1.042	1.471	0.2485
ABC	1.083	2	0.542	0.765	0.4867
Error	8.500	12	0.708		
Total	336.625	23			

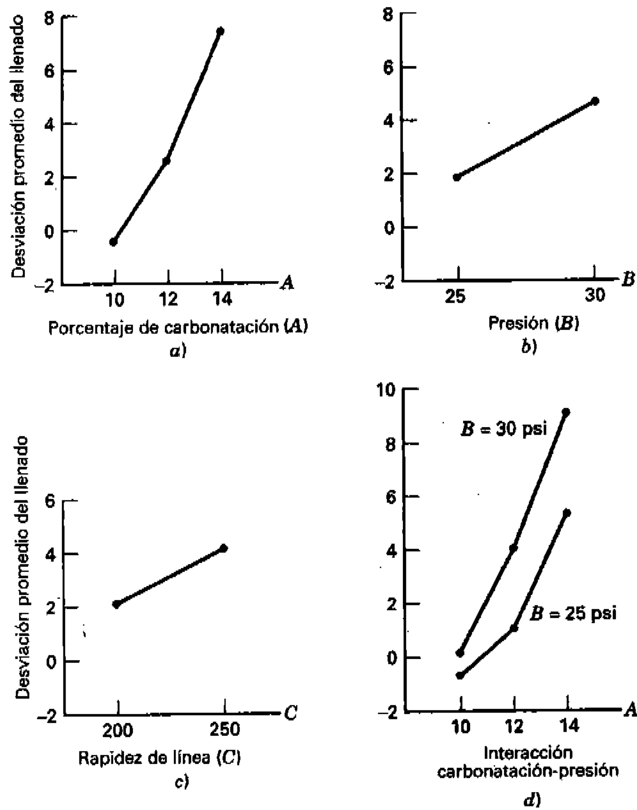


Figura 5-16 Gráficas de los efectos principales y la interacción del ejemplo 5-3. *a*) Porcentaje de carbonatación (*A*), *b*) presión (*B*), *c*) rapidez de línea (*C*), *d*) interacción carbonatación-presión.

promedio observada de la altura de llenado objetivo con los tres diferentes niveles de carbonatación para este conjunto de condiciones de operación. Ahora, el nivel de la carbonatación no puede actualmente controlarse perfectamente en el proceso de manufactura, y la distribución normal indicada con la línea continua de la figura 5-17 es una aproximación de la variabilidad de los niveles de carbonatación que se

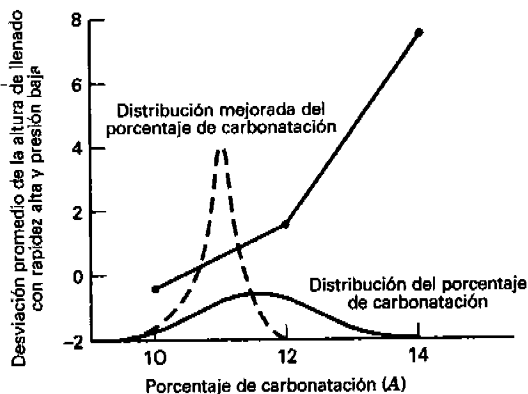


Figura 5-17 Desviación promedio de la altura de llenado con rapidez alta y presión baja para diferentes niveles de carbonatación.

registran actualmente. Como el proceso es impactado por los valores del nivel de carbonatación sacado de esta distribución, la fluctuación de las alturas de llenado será considerable. Esta variabilidad de las alturas de llenado podría reducirse si la distribución de los valores del nivel de carbonatación siguieran la distribución normal indicada con la línea punteada de la figura 5-17. La reducción de la desviación estándar de la distribución del nivel de carbonatación se consiguió finalmente mejorando el control de la temperatura durante la manufactura.

Se señaló ya que si todos los factores de un experimento factorial son fijos, la construcción del estadístico de prueba es directa. El estadístico para probar cualquier efecto principal o interacción se forma siempre dividiendo el cuadrado medio del efecto principal o la interacción por el cuadrado medio del error. Sin embargo, si el experimento factorial incluye uno o más **factores aleatorios,** la construcción del estadístico de prueba no siempre se hace de esta manera. Es necesario examinar los cuadrados medios esperados para determinar las pruebas correctas. La revisión completa de los experimentos con factores aleatorios se pospone hasta el capítulo 12.

5-5 AJUSTE DE CURVAS Y SUPERFICIES DE RESPUESTA

Se ha visto que puede resultar útil ajustar una **curva de respuesta** a los niveles de un factor cuantitativo para que el experimentador cuente con una ecuación que relacione la respuesta con el factor. Esta ecuación podría utilizarse para hacer interpolaciones, es decir, para predecir la respuesta en niveles intermedios entre los factores, respecto de los que se utilizaron realmente en el experimento. Cuando al menos dos de los factores son cuantitativos, puede ajustarse una **superficie de respuesta** para predecir y con varias combinaciones de los factores del diseño. En general, se usan **métodos de regresión lineal** para ajustar estos modelos a los datos experimentales. Este procedimiento se ilustra en la sección 3-5.1 para un experimento con un solo factor. A continuación se presentan dos ejemplos que incluyen experimentos factoriales. Se utilizará un paquete de software de computadora para generar los modelos de regresión. Para mayor información acerca del análisis de regresión, referirse al capítulo 10 y al material suplementario del texto de este capítulo.

EJEMPLO 5-4

Considere el experimento que se describe en el ejemplo 5-1. El factor temperatura es cuantitativo y el tipo de material es cualitativo. Además, hay tres niveles de la temperatura. Por consiguiente, puede calcularse un efecto de la temperatura lineal y uno cuadrático para estudiar la forma en que la temperatura afecta la vida de la batería. En la tabla 5-15 se presenta la salida condensada de *Design-Expert* para este experimento, donde se supone que la temperatura es cuantitativa y el tipo de material es cualitativo.

El análisis de varianza de la tabla 5-15 indica que la fuente de variabilidad "modelo" se ha subdividido en varios componentes. Los componentes "A" y "A^2" representan los efectos lineal y cuadrático de la temperatura, y "B" representa el efecto principal del factor tipo de material. Recuerde que el tipo de material es un factor cualitativo con tres niveles. Los términos "AB" y "A^2B" son las interacciones del factor temperatura lineal y cuadrático con el tipo de material.

Los valores P indican que A^2 y AB no son significativos, mientras que el término A^2B es significativo. Con frecuencia se piensa en eliminar los términos o factores no significativos del modelo, pero en este

Tabla 5-15 Salida de *Design-Expert* para el ejemplo 5-4

Response: Life in hr
ANOVA for Response Surface Reduced Cubic Model
Analysis of variance table [Partial sum of squares]

Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	59416.22	8	7427.03	11.00	<0.0001	significant
A	*39042.67*	*1*	*39042.67*	*57.82*	*<0.0001*	
B	*10683.72*	*2*	*5341.86*	*7.91*	*0.0020*	
A^2	*76.06*	*1*	*76.06*	*0.11*	*0.7398*	
AB	*2315.08*	*2*	*1157.54*	*1.71*	*0.1991*	
A^2B	*7298.69*	*2*	*3649.35*	*5.40*	*0.0106*	
Residual	18230.75	27	675.21			
Lack of Fit	*0.000*	*0*				
Pure Error	*18230.75*	*27*	*675.21*			
Cor Total	77646.97	35				

Std. Dev.	25.98	R-Squared	0.7652
Mean	105.53	Adj R-Squared	0.6956
C.V.	24.62	Pred R-Squared	0.5826
PRESS	32410.22	Adeq Precision	8.178

Term	Coefficient Estimate	DF	Standard Error	95% CI Low	95% CI High	VIF
Intercept	107.58	1	7.50	92.19	122.97	
A-Temp	−40.33	1	5.30	−51.22	−29.45	1.00
B[1]	−50.33	1	10.61	−72.10	−28.57	
B[2]	12.17	1	10.61	−9.60	33.93	
A^2	−3.08	1	9.19	−21.93	15.77	1.00
AB[1]	1.71	1	7.50	−13.68	17.10	
AB[2]	−12.79	1	7.50	−28.18	2.60	
A^2B[1]	41.96	1	12.99	15.30	68.62	
A^2B[2]	−14.04	1	12.99	−40.70	12.62	

Final Equation in Terms of Coded Factors:

Life =
+107.58
−40.33 *A
−50.33 *B[1]
+12.17 *B[2]
−3.08 *A^2
+1.71 *AB[1]
−12.79 *AB[2]
+41.96 *A^2B[1]
−14.04 *A^2B[2]

Final Equation in Terms of Actual Factors:

Material Type 1
Life =
+169.38017
−2.48860 *Temp
+0.012851 *Temp^2

Material Type 2
Life =
+159.62397
−0.17901 *Temp
+0.41627 *Temp^2

Material Type 3
Life =
+132.76240
+0.89264 *Temp
−0.43218 *Temp^2

caso eliminar A^2 y AB y conservar A^2B resultará en un modelo que no es **jerárquico.** El **principio de jerarquía** establece que si un modelo contiene un término de orden superior (tal como A^2B), deberá contener también todos los términos de orden inferior que lo componen (A^2 y AB en este caso). La jerarquía promueve un tipo de consistencia interna en un modelo, y muchos constructores de modelos estadísticos siguen rigurosamente este principio. Sin embargo, la jerarquía no es siempre una buena idea, y muchos modelos en realidad funcionan mejor como ecuaciones de predicción que no incluyen los términos no significativos que propone la jerarquía. Para mayor información, ver el material suplementario del texto de este capítulo.

La salida de computadora incluye también estimaciones de los coeficientes del modelo y una ecuación para la predicción final de la vida de la batería en términos de factores codificados. En esta ecuación, los niveles de la temperatura son $A = -1, 0, +1$, respectivamente, cuando la temperatura está en los niveles bajo, intermedio y alto (15, 70, 125°F). Las variables $B[1]$ y $B[2]$ son **variables indicadoras** codificadas que se definen de la siguiente manera:

	Tipo de material		
	1	2	3
$B[1]$	1	0	–1
$B[2]$	0	1	–1

Hay también ecuaciones para la predicción de la vida de la batería en términos de los niveles de los factores reales. Observe que como el tipo de material es un factor cualitativo, hay una ecuación para la vida predicha como una función de la temperatura para cada tipo de material. En la figura 5-18 se muestran

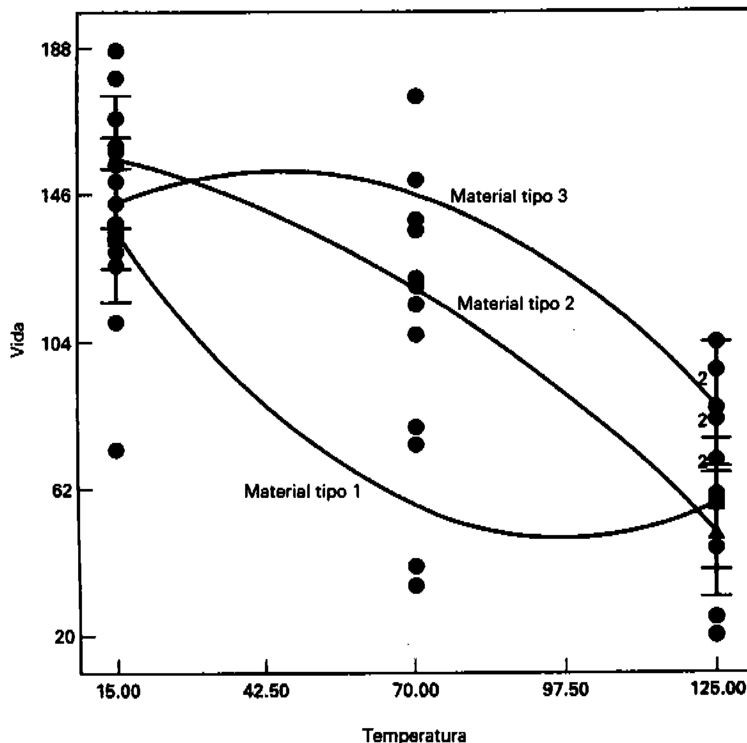


Figura 5-18 La vida predicha como una función de la temperatura para los tres tipos de material, ejemplo 5-4.

las curvas de respuesta generadas por estas tres ecuaciones de predicción. Compárense con la gráfica de la interacción de dos factores para este experimento de la figura 5-9.

Si varios de los factores de un experimento factorial son cuantitativos, puede usarse una **superficie de respuesta** para modelar la relación entre y y los factores del diseño. Además, los efectos de los factores cuantitativos pueden representarse con efectos polinomiales con un solo grado de libertad. De manera similar, es posible hacer la partición de las interacciones de factores cuantitativos en componentes de interacción con un solo grado de libertad. Esto se ilustra en el ejemplo siguiente.

EJEMPLO 5-5

Se piensa que la vida efectiva de una herramienta de corte instalada en una máquina controlada numéricamente se afecta por la velocidad de corte y el ángulo de la herramienta. Se seleccionan tres velocidades y tres ángulos, y se lleva a cabo un experimento factorial con dos réplicas. En la tabla 5-16 se muestran los datos codificados. Los números de las celdas encerrados en círculos son los totales de las celdas $\{y_{ij.}\}$.

En la tabla 5-17 se presenta la salida condensada de *Design-Expert* para este ejemplo. Los términos A y A^2 son los efectos lineal y cuadrático del ángulo de la herramienta, y B y B^2 son los efectos lineal y cuadrático de la velocidad. Los términos AB, A^2B, AB^2 y A^2B^2 representan los **componentes** lineal × lineal, cuadrático × lineal, lineal × cuadrático y cuadrático × cuadrático de la interacción de dos factores. Aun cuando hay algunos valores P grandes, se han conservado todos los términos del modelo para respetar la jerarquía. En la ecuación de predicción expresada en factores codificados se utilizan los niveles −1, 0 y +1 de A y B para representar los niveles bajo, intermedio y alto, respectivamente, de estos factores.

En la figura 5-19 se presenta la **gráfica de contorno** de la superficie generada por la ecuación de predicción de la vida de la herramienta. El examen de esta **superficie de respuesta** indica que la vida máxima de la herramienta se consigue con velocidades de corte de alrededor de 150 rpm y ángulos de la herramienta de 25°. La gráfica de la superficie de respuesta tridimensional de la figura 5-20 proporciona en esencia la misma información, pero ofrece una perspectiva diferente, y en ocasiones más útil, de la superficie de respuesta de la vida de la herramienta. La exploración de las superficies de respuesta es un aspecto muy importante del diseño experimental, el cual se estudiará en detalle en el capítulo 11.

Tabla 5-16 Datos del experimento de la vida de la herramienta de corte

Ángulo de la herramienta (grados)	Velocidad de corte (pulg/min)						
	125		150		175		$y_{i..}$
15	−2 −1	(−3)	−3 0	(−3)	2 3	(5)	−1
20	0 2	(2)	1 3	(4)	4 6	(10)	16
25	−1 0	(−1)	5 6	(11)	0 −1	(−1)	9
$y_{.j.}$	−2		12		14		$24 = y_{...}$

Tabla 5-17 Salida de *Design-Expert* para el ejemplo 5-5

Response: Life in Hours
ANOVA for Response Surface Reduced Order 4 Model
Analysis of variance table [Partial sum of squares]

Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	111.00	8	13.87	9.61	0.0013	significant
A	*49.00*	*1*	*49.00*	*33.92*	*0.0003*	
B	*16.00*	*1*	*16.00*	*11.08*	*0.0088*	
A^2	*0.000*	*1*	*0.000*	*0.000*	*1.0000*	
B^2	*1.33*	*1*	*1.33*	*0.92*	*0.3618*	
AB	*8.00*	*1*	*8.00*	*5.54*	*0.0431*	
A^2B	*2.67*	*1*	*2.67*	*1.85*	*0.2073*	
AB^2	*42.67*	*1*	*42.67*	*29.54*	*0.0004*	
A^2B^2	*8.00*	*1*	*8.00*	*5.54*	*0.0431*	
Residual	13.00	9	1.44			
Lack of Fit	*0.000*	*0*				
Pure Error	*13.00*	*9*	*1.44*			
Cor Total	124.00	17				

Std. Dev.	1.20	R-Squared	0.8952
Mean	1.33	Adj R-Squared	0.8020
C.V.	90.14	Pred R-Squared	0.5806
PRESS	52.00	Adeq Precision	8.237

Factor	Coefficient Estimate	DF	Standard Error	95% CI Low	95% CI High	VIF
Intercept	2.00	1	0.85	0.078	3.92	
A-Tool Angle	3.50	1	0.60	2.14	4.86	3.00
B-Speed	2.00	1	0.60	0.64	3.36	3.00
A^2	0.000	1	1.04	−2.35	2.35	3.00
B^2	1.00	1	1.04	−1.35	3.35	3.00
AB	−1.00	1	0.42	−1.96	−0.039	1.00
A^2B	−1.00	1	0.74	−2.66	0.66	3.00
AB^2	−4.00	1	0.74	−5.66	−2.34	3.00
A^2B^2	−3.00	1	1.27	−5.88	−0.12	5.00

Final Equation in Terms of Coded Factors:

Life =
+2.00
+3.50 *A
+2.00 *B
+0.000 *A^2
+1.00 *B^2
−1.00 *A *B
−1.00 *A^2 *B
−4.00 *A *B^2
−3.00 *A^2 *B^2

Final Equation in Terms of Actual Factors:

Life =
−1068.00000
+136.30000 *Tool Angle
+14.48000 *Speed
−4.08000 *Tool $Angle^2$
−0.049600 *$Speed^2$
−1.86400 *Tool Angle *Speed
+0.056000 *Tool $Angle^2$ *Speed
+6.40000E−003 *Tool Angle *$Speed^2$
−1.92000E−004 *Tool $Angle^2$ *$Speed^2$

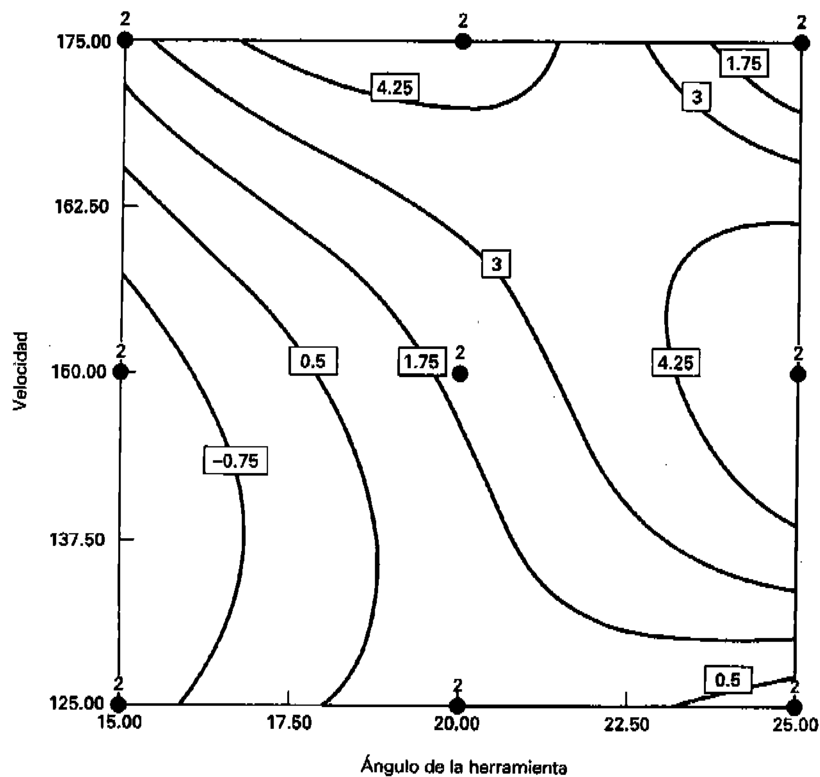


Figura 5-19 Gráfica de contorno bidimensional de la superficie de respuesta de la vida de la herramienta del ejemplo 5-5.

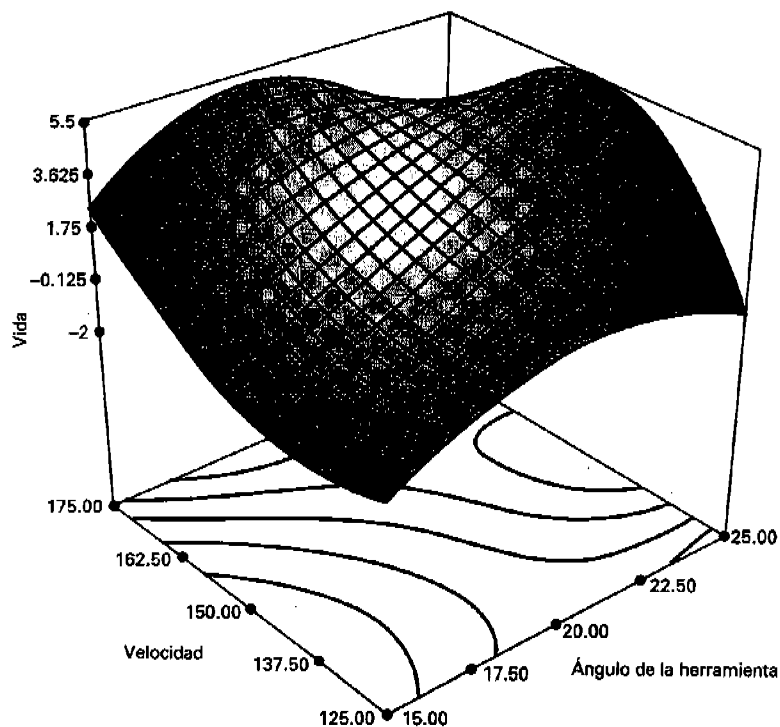


Figura 5-20 Superficie de respuesta tridimensional de la vida de la herramienta del ejemplo 5-5.

5-6 FORMACIÓN DE BLOQUES EN UN DISEÑO FACTORIAL

Se han revisado los diseños factoriales en el contexto de un experimento **completamente aleatorizado.** En ocasiones no es factible o práctico hacer la aleatorización completa de todas las corridas de un diseño factorial. Por ejemplo, la presencia de un factor perturbador puede hacer necesario que el experimento se corra en **bloques.** Los conceptos básicos de la formación de bloques se analizaron en el capítulo 4 en el contexto de un experimento con un solo factor. Ahora se indica la forma en que la formación de bloques puede incorporarse en un diseño factorial. Otros aspectos de la formación de bloques en diseños factoriales se presentan en los capítulos 7, 8, 9 y 13.

Considere un experimento factorial con dos factores (A y B) y n réplicas. El modelo estadístico lineal de este diseño es

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \varepsilon_{ijk} \quad \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, b \\ k = 1, 2, \ldots, n \end{cases} \tag{5-36}$$

donde τ_i, β_j y $(\tau\beta)_{ij}$ representan los efectos de los factores A, B y la interacción AB, respectivamente. Suponga ahora que para realizar este experimento se necesita una materia prima particular. Esta materia prima está disponible en lotes cuyo tamaño no es suficiente para permitir que se corran *todas* las abn combinaciones de los tratamientos con el *mismo* lote. Sin embargo, si un lote contiene material suficiente para hacer ab observaciones, entonces un diseño alternativo es correr cada una de las n réplicas utilizando un lote separado de materia prima. Por consiguiente, los lotes de materia prima representan una restricción sobre la aleatorización o un **bloque,** y se corre una sola réplica de un experimento factorial completo dentro de cada bloque. El modelo de los efectos para este nuevo diseño es

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \delta_k + \varepsilon_{ijk} \quad \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, b \\ k = 1, 2, \ldots, n \end{cases} \tag{5-37}$$

donde δ_k es el efecto del bloque k-ésimo. Desde luego, dentro de un bloque el orden en que se corren las combinaciones de los tratamientos está completamente aleatorizado.

En el modelo (ecuación 5-37) se supone que la interacción entre los bloques y los tratamientos es insignificante. Anteriormente se estableció el mismo supuesto en el análisis de diseños de bloques aleatorizados. Si estas interacciones existen, no pueden separarse del componente del error. De hecho, el término del error en este modelo se compone en realidad de las interacciones $(\tau\delta)_{ik}$, $(\beta\delta)_{jk}$ y $(\tau\beta\delta)_{ijk}$. En la tabla 5-18 se describe el análisis de varianza. La disposición tiene un gran parecido con la de un diseño factorial, con la suma de cuadrados del error reducida por la suma de cuadrados de los bloques. En lo que a los cálculos se refiere, la suma de cuadrados de los bloques se encuentra como la suma de cuadrados entre los totales de los n bloques $\{y_{..k}\}$.

En el ejemplo anterior, la aleatorización se restringió al interior de un lote de materia prima. En la práctica, una diversidad de fenómenos pueden producir restricciones sobre la aleatorización, como el tiempo, los operadores, etc. Por ejemplo, si el experimento factorial completo no pudo correrse en un día, entonces el experimentador podría correr una réplica completa el día 1, una segunda réplica el día 2, etc. Por consiguiente, cada día sería un bloque.

EJEMPLO 5-6 ··

Un ingeniero estudia los métodos para mejorar la capacidad para detectar objetivos en el campo de acción de un radar. Dos factores que el ingeniero considera importantes son la cantidad de ruido de fondo, o "desorden de terreno", en el campo de acción del radar y el tipo de filtro colocado sobre la pantalla. Se

Tabla 5-18 Análisis de varianza de un diseño factorial de dos factores en bloques completos aleatorizados

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio esperado	F_0
Bloques	$\frac{1}{ab}\sum_k y_{..k}^2 - \frac{y_{...}^2}{abn}$	$n-1$	$\sigma^2 + ab\sigma_\delta^2$	
A	$\frac{1}{bn}\sum_i y_{i..}^2 - \frac{y_{...}^2}{abn}$	$a-1$	$\sigma^2 + \frac{bn\sum \tau_i^2}{a-1}$	$\frac{MS_A}{MS_E}$
B	$\frac{1}{an}\sum_j y_{.j.}^2 - \frac{y_{...}^2}{abn}$	$b-1$	$\sigma^2 + \frac{an\sum \beta_j^2}{b-1}$	$\frac{MS_B}{MS_E}$
AB	$\frac{1}{n}\sum_i\sum_j y_{ij.}^2 - \frac{y_{...}^2}{abn} - SS_A - SS_B$	$(a-1)(b-1)$	$\sigma^2 + \frac{n\sum\sum(\tau\beta)_{ij}^2}{(a-1)(b-1)}$	$\frac{MS_{AB}}{MS_E}$
Error	Sustracción	$(ab-1)(n-1)$	σ^2	
Total	$\sum_i\sum_j\sum_k y_{ijk}^2 - \frac{y_{...}^2}{abn}$	$abn-1$		

diseña un experimento utilizando tres niveles del desorden de terreno y dos tipos de filtro. Estos factores se considerarán fijos. El experimento se lleva a cabo seleccionando al azar una combinación de los tratamientos (nivel del desorden de terreno y tipo de filtro) e introduciendo después una señal que representa el objetivo en el campo de acción del radar. La intensidad de este objetivo se incrementa hasta que el operador lo observa. Entonces se mide el nivel de intensidad en el momento de la detección como la variable de respuesta. Debido a la disponibilidad de los operadores, es conveniente seleccionar un operador y mantenerlo en el sistema hasta que se han realizado todas las corridas necesarias. Además, los operadores difieren en su habilidad y capacidad para operar el sistema. Por consiguiente, parece lógico usar los operadores como bloques. Se seleccionan cuatro operadores al azar. Una vez que se ha elegido a un operador, el orden en que se corren las seis combinaciones de los tratamientos se determina aleatoriamente. Por lo tanto, se tiene una corrida de un experimento factorial 3×2 en un bloque completo aleatorizado. Los datos se presentan en la tabla 5-19.

El modelo lineal para este experimento es

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \delta_k + \varepsilon_{ijk} \quad \begin{cases} i = 1, 2, 3 \\ j = 1, 2 \\ k = 1, 2, 3, 4 \end{cases}$$

donde τ_i representa el efecto del desorden de terreno, β_j representa el efecto del tipo de filtro, $(\tau\beta)_{ij}$ es la interacción, δ_k es el efecto del bloque y ε_{ijk} es el componente NID$(0, \sigma^2)$ del error. Las sumas de cuadrados del desorden de terreno, del tipo de filtro y de su interacción se calculan de la manera usual. La suma de

Tabla 5-19 Nivel de intensidad al detectarse el objetivo

Operadores (bloques)	1		2		3		4	
Tipo de filtro	1	2	1	2	1	2	1	2
Desorden de terreno								
Bajo	90	86	96	84	100	92	92	81
Intermedio	102	87	106	90	105	97	96	80
Alto	114	93	112	91	108	95	98	83

Tabla 5-20 Análisis de varianza del ejemplo 5-6

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0	Valor P
Desorden de terreno (G)	335.58	2	167.79	15.13	0.0003
Tipo de filtro (F)	1066.67	1	1066.67	96.19	<0.0001
GF	77.08	2	38.54	3.48	0.0573
Bloques	402.17	3	134.06		
Error	166.33	15	11.09		
Total	2047.83	23			

cuadrados debida a los bloques se encuentra a partir de los totales de los operadores $\{y_{..k}\}$ de la siguiente manera:

$$SS_{\text{Bloques}} = \frac{1}{ab}\sum_{k=1}^{n} y_{..k}^2 - \frac{y_{...}^2}{abn}$$

$$= \frac{1}{(3)(2)}[(572)^2 + (579)^2 + (597)^2 + (530)^2] - \frac{(2278)^2}{(3)(2)(4)}$$

$$= 402.17$$

En la tabla 5-20 se resume el análisis de varianza completo de este experimento. La presentación de la tabla 5-20 indica que todos los efectos se probaron dividiendo sus cuadrados medios por el cuadrado medio del error. Tanto el desorden de terreno como el tipo de filtro son significativos en el nivel de 1%, mientras que su interacción sólo es significativa en el nivel de 10%. Por lo tanto, se concluye que tanto el nivel del desorden de terreno como el tipo de filtro de campo usado en la pantalla afectan la habilidad del operador para detectar el objetivo, y existe cierta evidencia de una ligera interacción entre ambos factores.

En el caso de dos restricciones sobre la aleatorización, cada una con p niveles, si el número de combinaciones de los tratamientos en un diseño factorial de k factores es exactamente igual al número de niveles de la restricción, es decir, si $p = ab \ldots m$, entonces el diseño factorial puede correrse en un cuadrado latino $p \times p$. Por ejemplo, considere una modificación del experimento de la detección del objetivo en el radar del ejemplo 5-6. Los factores de este experimento son el tipo de filtro (dos niveles) y el desorden de terreno (tres niveles), y los operadores se consideran como bloques. Suponga ahora que debido a limitaciones de tiempo, sólo pueden hacerse seis corridas por día. Por lo tanto, los días se convierten en una segunda restricción sobre la aleatorización, lo cual resulta en un diseño del cuadrado latino 6 × 6, como se muestra en la tabla 5-21. En esta tabla se han usado las letras minúsculas f_i y g_j para representar los niveles i-ésimo y j-ésimo del tipo de filtro y del desorden de terreno, respectivamente. Es decir, f_1g_2 representa el filtro tipo 1 y un desorden de terreno intermedio. Observe que se necesitan ahora seis operadores, en lugar de los cuatro del experimento original, por lo que el número de combinaciones de tratamientos en el diseño factorial 3 × 2 es exactamente igual al número de niveles de restricción. Además, en este diseño cada operador se usaría una sola vez en cada día. Las letras latinas A, B, C, D, E y F representan las 3 × 2 = 6 combinaciones de tratamientos del diseño factorial como sigue: $A = f_1g_1$, $B = f_1g_2$, $C = f_1g_3$, $D = f_2g_1$, $E = f_2g_2$ y $F = f_2g_3$.

Tabla 5-21 El experimento de la detección del radar realizado en un cuadrado latino 6 × 6

	Operador					
Día	1	2	3	4	5	6
1	$A(f_1g_1 = 90)$	$B(f_1g_2 = 106)$	$C(f_1g_3 = 108)$	$D(f_2g_1 = 81)$	$F(f_2g_3 = 90)$	$E(f_2g_2 = 88)$
2	$C(f_1g_3 = 114)$	$A(f_1g_1 = 96)$	$B(f_1g_2 = 105)$	$F(f_2g_3 = 83)$	$E(f_2g_2 = 86)$	$D(f_2g_1 = 84)$
3	$B(f_1g_2 = 102)$	$E(f_2g_2 = 90)$	$F(f_2g_3 = 95)$	$A(f_1g_1 = 92)$	$D(f_2g_1 = 85)$	$C(f_1g_3 = 104)$
4	$E(f_2g_2 = 887)$	$D(f_2g_1 = 84)$	$A(f_1g_1 = 100)$	$B(f_1g_2 = 96)$	$C(f_1g_3 = 110)$	$F(f_2g_3 = 91)$
5	$F(f_2g_3 = 93)$	$C(f_1g_3 = 112)$	$D(f_2g_1 = 92)$	$E(f_2g_2 = 80)$	$A(f_1g_1 = 90)$	$B(f_1g_2 = 98)$
6	$D(f_2g_1 = 86)$	$F(f_2g_3 = 91)$	$E(f_2g_2 = 97)$	$C(f_1g_3 = 98)$	$B(f_1g_2 = 100)$	$A(f_1g_1 = 92)$

Los cinco grados de libertad entre las seis letras latinas corresponden a los efectos principales del tipo de filtro (un grado de libertad), el desorden de terreno (dos grados de libertad) y su interacción (dos grados de libertad). El modelo estadístico lineal de este diseño es

$$y_{ijkl} = \mu + \alpha_i + \tau_j + \beta_k + (\tau\beta)_{jk} + \theta_l + \varepsilon_{ijkl} \quad \begin{cases} i = 1, 2, \ldots, 6 \\ j = 1, 2, 3 \\ k = 1, 2 \\ l = 1, 2, \ldots, 6 \end{cases} \tag{5-38}$$

donde τ_j y β_k son los efectos del desorden de terreno y del tipo de filtro, respectivamente, y α_i y θ_l representan las restricciones sobre la aleatorización de los días y los operadores, respectivamente. Para calcular las sumas de cuadrados, la siguiente tabla de dos vías de los totales de los tratamientos es útil:

Desorden de terreno	Filtro tipo 1	Filtro tipo 2	$y_{.j..}$
Bajo	560	512	1072
Intermedio	607	528	1135
Alto	646	543	1189
$y_{..k.}$	1813	1583	$3396 = y_{....}$

Además, los totales de los renglones y las columnas son

Renglones ($y_{.jkl}$):	563	568	568	568	565	564
Columnas ($y_{ijk.}$):	572	579	597	530	561	557

En la tabla 5-22 se resume el análisis de varianza. Se ha agregado una columna a esta tabla que indica cómo se determina el número de grados de libertad de cada suma de cuadrados.

Tabla 5-22 Análisis de varianza del experimento de la detección en el radar realizado como un diseño factorial 3 × 2 en un cuadrado latino

Fuente de variación	Suma de cuadrados	Grados de libertad	Fórmula general para los grados de libertad	Cuadrado medio	F_0	Valor *P*
Desorden de terreno, *G*	571.50	2	$a - 1$	285.75	28.86	<0.0001
Tipo de filtro, *F*	1469.44	1	$b - 1$	1469.44	148.43	<0.0001
GF	126.73	2	$(a-1)(b-1)$	63.37	6.40	0.0071
Días (renglones)	4.33	5	$ab - 1$	0.87		
Operadores (columnas)	428.00	5	$ab - 1$	85.60		
Error	198.00	20	$(ab-1)(ab-2)$	9.90		
Total	2798.00	36	$(ab)^2 - 1$			

5-7 PROBLEMAS

5-1. Se estudia el rendimiento de un proceso químico. Se piensa que las dos variables más importantes son la presión y la temperatura. Se seleccionan tres niveles de cada factor y se lleva a cabo un experimento factorial con dos réplicas. Los datos del rendimiento son:

	Presión (psig)		
Temperatura (°C)	200	215	230
150	90.4	90.7	90.2
	90.2	90.6	90.4
160	90.1	90.5	89.9
	90.3	90.6	90.1
170	90.5	90.8	90.4
	90.7	90.9	90.1

a) Analizar los datos y sacar conclusiones. Utilizar $\alpha = 0.05$.
b) Construir las gráficas de los residuales apropiadas y comentar la adecuación del modelo.
c) ¿Bajo qué condiciones debería operarse este proceso?

5-2. Un ingeniero sospecha que el acabado superficial de una pieza metálica se afecta por la velocidad de alimentación y la profundidad de corte. Selecciona tres velocidades de alimentación y cuatro profundidades de corte. Después realiza un experimento factorial y obtiene los siguientes datos:

Velocidad de alimentación	Profundiad de corte (pulg)			
(pulg/min)	0.15	0.18	0.20	0.25
	74	79	82	99
0.20	64	68	88	104
	60	73	92	96
	92	98	99	104
0.25	86	104	108	110
	88	88	95	99
	99	104	108	114
0.30	98	99	110	111
	102	95	99	107

a) Analizar los datos y sacar conclusiones. Utilizar $\alpha = 0.05$.
b) Construir las gráficas de los residuales apropiadas y comentar la adecuación del modelo.
c) Obtener estimaciones puntuales del acabado superficial promedio con cada velocidad de alimentación.
d) Encontrar los valores P para las pruebas del inciso *a*.

5-3. Para los datos del problema 5-2, calcular la estimación de un intervalo de confianza de 95% de la diferencia media en la respuesta para velocidades de alimentación de 0.20 y 0.25 pulg/min.

5-4. En un artículo de *Industrial Quality Control* se describe un experimento para investigar el efecto del tipo de cristal y del tipo de fósforo sobre la brillantez de un cinescopio. La variable de respuesta es la corriente (en microamperes) necesaria para obtener un nivel de brillantez específico. Los datos son los siguientes:

Tipo de cristal	Tipo de fósforo		
	1	2	3
1	280	300	290
	290	310	285
	285	295	290
2	230	260	220
	235	240	225
	240	235	230

a) ¿Existe algún indicio de que alguno de los dos factores influye en la brillantez? Utilizar $\alpha = 0.05$.
b) ¿Los dos factores interactúan? Utilizar $\alpha = 0.05$.
c) Analizar los residuales de este experimento.

5-5. Johnson y Leone (*Statistics and Experimental Design in Engineering and the Physical Sciences*, John Wiley) describen un experimento realizado para investigar la torcedura de placas de cobre. Los dos factores estudiados fueron la temperatura y el contenido de cobre de las placas. La variable de respuesta fue una medida de la cantidad de torcedura. Los datos fueron los siguientes:

Temperatura (°C)	Contenido de cobre (%)			
	40	60	80	100
50	17, 20	16, 21	24, 22	28, 27
75	12, 9	18, 13	17, 12	27, 31
100	16, 12	18, 21	25, 23	30, 23
125	21, 17	23, 21	23, 22	29, 31

a) ¿Existe algún indicio de que alguno de los dos factores afecta la cantidad de torcedura? ¿Hay alguna interacción entre los factores? Utilizar $\alpha = 0.05$.
b) Analizar los residuales de este experimento.
c) Graficar la torcedura promedio con cada nivel del contenido de cobre y compararlas con una distribución t con la escala apropiada. Describir las diferencias en los efectos de los diversos niveles del contenido de cobre sobre la torcedura. Si es deseable una torcedura baja, ¿qué nivel del contenido de cobre debería especificarse?
d) Suponga que no es sencillo controlar la temperatura en el medio ambiente donde van a usarse las placas de cobre. ¿Este hecho modifica la respuesta que se dio para el inciso *c*?

5-6. Se estudian los factores que influyen en la resistencia a la ruptura de una fibra sintética. Se eligen cuatro máquinas de producción y tres operadores y se corre un experimento factorial utilizando fibra del mismo lote de producción. Los resultados son los siguientes:

Operador	Máquina			
	1	2	3	4
1	109	110	108	110
	110	115	109	108
2	110	110	111	114
	112	111	109	112
3	116	112	114	120
	114	115	119	117

a) Analizar los datos y sacar conclusiones. Utilizar $\alpha = 0.05$.

b) Construir las gráficas de los residuales apropiadas y comentar la adecuación del modelo.

5-7. Un ingeniero mecánico estudia la fuerza de empuje desarrollada por una taladradora. Sospecha que la velocidad de taladrado y la velocidad de alimentación del material son los factores más importantes. Selecciona cuatro velocidades de alimentación y usa una velocidad de taladrado alta y otra baja elegidas para representar las condiciones de operación extremas. Obtiene los siguientes resultados. Analizar los datos y sacar conclusiones. Utilizar $\alpha = 0.05$.

	Velocidad de alimentación			
Velocidad de taladrado	0.015	0.030	0.045	0.060
125	2.70	2.45	2.60	2.75
	2.78	2.49	2.72	2.86
200	2.83	2.85	2.86	2.94
	2.86	2.80	2.87	2.88

5-8. Se realiza un experimento para estudiar la influencia de la temperatura de operación y tres tipos de placas de recubrimiento de cristal, en la salida luminosa de un tubo de osciloscopio. Se registraron los siguientes datos:

Tipo de cristal	Temperatura		
	100	125	150
	580	1090	1392
1	568	1087	1380
	570	1085	1386
	550	1070	1328
2	530	1035	1312
	579	1000	1299
	546	1045	867
3	575	1053	904
	599	1066	889

a) Utilizar $\alpha = 0.05$ en el análisis. ¿Existe un efecto de interacción significativo? ¿El tipo de cristal o la temperatura afectan la respuesta? ¿A qué conclusiones se llega?

b) Ajustar un modelo apropiado que relacione la salida luminosa con el tipo de cristal y la temperatura.

c) Analizar los residuales de este experimento. Comentar la adecuación de los modelos que se hayan considerado.

5-9. Considere el experimento del problema 5-1. Ajustar un modelo apropiado a los datos de la respuesta. Usar este modelo como guía para las condiciones de operación del proceso.

5-10. Usar la prueba de Tukey para determinar los niveles del factor presión que son significativamente diferentes para los datos del problema 5-1.

5-11. Se llevó a cabo un experimento para determinar si la temperatura de cocción o la posición en el horno afectaban el espesor del endurecimiento de un ánodo de carbono. Los datos se presentan a continuación:

	Temperatura (°C)		
Posición	800	825	850
1	570	1063	565
	565	1080	510
	583	1043	590
2	528	988	526
	547	1026	538
	521	1004	532

Suponga que se considera que no existe ninguna interacción. Desarrollar el modelo estadístico. Realizar el análisis de varianza y probar las hipótesis sobre los efectos principales. ¿Qué conclusiones pueden sacarse? Comentar la adecuación del modelo.

5-12. Deducir los cuadrados medios esperados para un análisis de varianza de dos factores con una observación por celda, suponiendo que ambos factores son fijos.

5-13. Considere los siguientes datos de un experimento factorial de dos factores. Analizar los datos y sacar conclusiones. Realizar una prueba de no aditividad. Utilizar $\alpha = 0.05$.

	Factor de la columna			
Factor del renglón	1	2	3	4
1	36	39	36	32
2	18	20	22	20
3	30	37	33	34

5-14. Se piensa que la resistencia al corte de un adhesivo se afecta por la presión de aplicación y la temperatura. Se realiza un experimento factorial en el que ambos factores se suponen fijos. Analizar los datos y sacar conclusiones. Realizar una prueba de no aditividad.

Presión	Temperatura (°F)		
(lb/pulg^2)	250	260	270
120	9.60	11.28	9.00
130	9.69	10.10	9.57
140	8.43	11.01	9.03
150	9.98	10.44	9.80

5-15. Considere el modelo de tres factores

$$y_{ijk} = \mu + \tau_i + \beta_j + \gamma_k + (\tau\beta)_{ij} + (\beta\gamma)_{jk} + \varepsilon_{ijk} \quad \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, b \\ k = 1, 2, \ldots, c \end{cases}$$

Observe que hay una sola réplica. Suponiendo que los tres factores son fijos, desarrollar la tabla del análisis de varianza, incluyendo los cuadrados medios esperados. ¿Qué se usaría como "error experimental" para probar las hipótesis?

5-16. El porcentaje de la concentración de madera dura en la pulpa bruta, la presión de la cuba y el tiempo de cocción de la pulpa se investigan en cuanto a sus efectos sobre la resistencia del papel. Se seleccionan tres niveles de la concentración de madera dura, tres niveles de la presión y dos tiempos de cocción. Se lleva a cabo un experimento factorial con dos réplicas, obteniéndose los siguientes datos:

Porcentaje de la concentración de madera dura	Tiempo de cocción 3.0 horas			Tiempo de cocción 4.0 horas		
	Presión			Presión		
	400	500	650	400	500	650
2	196.6	197.7	199.8	198.4	199.6	200.6
	196.0	196.0	199.4	198.6	200.4	200.9
4	198.5	196.0	198.4	197.5	198.7	199.6
	197.2	196.9	197.6	198.1	198.0	199.0
8	197.5	195.6	197.4	197.6	197.0	198.5
	196.6	196.2	198.1	198.4	197.8	199.8

a) Analizar los datos y sacar conclusiones. Utilizar $\alpha = 0.05$.

b) Construir las gráficas de los residuales apropiadas y comentar la adecuación del modelo.

c) ¿Bajo qué conjunto de condiciones debería operarse este proceso? ¿Por qué?

5-17. El departamento de control de calidad de una planta de acabados textiles estudia el efecto de varios factores sobre el teñido de una tela de algodón y fibras sintéticas utilizada para fabricar camisas para caballero. Se seleccionaron tres operadores, tres duraciones del ciclo y dos temperaturas, y se tiñeron tres ejemplares pequeños de la tela bajo cada conjunto de condiciones. La tela terminada se comparó con un patrón, y se le asignó una evaluación numérica. Los datos se presentan enseguida. Analizar los datos y sacar conclusiones. Comentar la adecuación del modelo.

	Temperatura					
	300°			350°		
	Operador			Operador		
Duración del ciclo	1	2	3	1	2	3
	23	27	31	24	38	34
40	24	28	32	23	36	36
	25	26	29	28	35	39
	36	34	33	37	34	34
50	35	38	34	39	38	36
	36	39	35	35	36	31
	28	35	26	26	36	28
60	24	35	27	29	37	26
	27	34	25	25	34	24

5-18. Suponga que en el problema 5-1 quiere rechazarse la hipótesis nula con una alta probabilidad si la diferencia entre el verdadero rendimiento promedio con dos presiones cualesquiera es mayor que 0.5. Si una estimación previa razonable de la desviación estándar del rendimiento es 0.1, ¿cuántas réplicas deberán correrse?

5-19. Se estudia el rendimiento de un proceso químico. Los dos factores de interés son la temperatura y la presión. Se seleccionan tres niveles de cada factor; sin embargo, sólo es posible hacer nueve corridas en un día. El ex-

perimentador corre una réplica completa en cada día. Los datos se muestran en la tabla siguiente. Analizar los datos, suponiendo que los días son bloques.

	Día 1 Presión			Día 2 Presión		
Temperatura	250	260	270	250	260	270
Baja	86.3	84.0	85.8	86.1	85.2	87.3
Intermedia	88.5	87.3	89.0	89.4	89.9	90.3
Alta	89.1	90.2	91.3	91.7	93.2	93.7

5-20. Considere los datos del problema 5-5. Analizar los datos, suponiendo que las réplicas son bloques.

5-21. Considere los datos del problema 5-6. Analizar los datos, suponiendo que las réplicas son bloques.

5-22. En un artículo de *Journal of Testing and Evaluation* (vol. 16, no. 2, pp. 508-515) se investigaron los efectos de la frecuencia de carga cíclica y de las condiciones ambientales sobre el crecimiento de las fisuras por fatiga con un esfuerzo constante de 22 MPa para un material particular. Los datos del experimento se presentan abajo (la respuesta es el índice de crecimiento de las fisuras por fatiga):

	Medio ambiente		
Frecuencia	Aire	H_2O	H_2O salada
10	2.29	2.06	1.90
	2.47	2.05	1.93
	2.48	2.23	1.75
	2.12	2.03	2.06
1	2.65	3.20	3.10
	2.68	3.18	3.24
	2.06	3.96	3.98
	2.38	3.64	3.24
0.1	2.24	11.00	9.96
	2.71	11.00	10.01
	2.81	9.06	9.36
	2.08	11.30	10.40

a) Analizar los datos de este experimento (utilizar $\alpha = 0.05$).

b) Analizar los residuales.

c) Repetir los análisis de los incisos *a* y *b* utilizando ln(*y*) como la respuesta. Comentar los resultados.

5-23. En un artículo de *IEEE Transactions on Electron Devices* se describe un estudio sobre el dopado del polisilicio. El experimento que se muestra a continuación es una variante de dicho estudio. La variable de respuesta es la corriente fundamental.

Dopado del polisilicio (iones)	Temperatura de revenido (°C) 900	950	1000
1×10^{20}	4.60	10.15	11.01
	4.40	10.20	10.58
2×10^{20}	3.20	9.38	10.81
	3.50	10.02	10.60

a) ¿Existe evidencia (con $\alpha = 0.05$) que indique que el nivel de dopado del polisilicio o la temperatura de fijación afecten la corriente fundamental?
b) Construir representaciones gráficas como ayuda para interpretar este experimento.
c) Analizar los residuales y comentar la adecuación del modelo.
d) ¿El modelo

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{22} x_2^2 + \beta_{12} x_1 x_2 + \varepsilon$$

está apoyado por este experimento (x_1 = nivel de dopado, x_2 = temperatura)? Estimar los parámetros de este modelo y graficar la superficie de respuesta.

6 Diseño factorial 2^k

6-1 INTRODUCCIÓN

Los diseños factoriales se usan ampliamente en experimentos que incluyen varios factores cuando es necesario estudiar el efecto conjunto de los factores sobre una respuesta. En el capítulo 5 se presentaron los métodos generales para el análisis de los diseños factoriales. Sin embargo, hay varios casos especiales del diseño factorial general que son importantes debido a su uso generalizado en el trabajo de investigación y porque constituyen las bases de otros diseños de gran valor práctico.

El más importante de estos casos especiales es el de k factores, cada uno con sólo dos niveles. Estos niveles pueden ser cuantitativos, como dos valores de temperatura, presión o tiempo, o bien cualitativos, como dos máquinas, dos operadores, los niveles "alto" y "bajo" de un factor, o quizá la presencia o ausencia de un factor. Una réplica completa de este diseño requiere $2 \times 2 \times \ldots \times 2 = 2^k$ observaciones y se le llama **diseño factorial 2^k.**

Este capítulo se enfoca en esta clase en extremo importante de diseños. A lo largo del capítulo se supone que 1) los factores son fijos, 2) los diseños son completamente aleatorizados y 3) se satisfacen los supuestos de normalidad usuales.

El diseño 2^k es de particular utilidad en las etapas iniciales del trabajo experimental, cuando probablemente se estén investigando muchos factores. Este diseño proporciona el menor número de corridas con las que pueden estudiarse k factores en un diseño factorial completo. Por consiguiente, estos diseños se usan ampliamente en los **experimentos de tamizado** o **selección de factores**.

Puesto que sólo hay dos niveles para cada factor, se supone que la respuesta es aproximadamente lineal en el rango elegido para los niveles de los factores. En muchos experimentos de tamizado de factores, cuando se acaba de iniciar el estudio del proceso o sistema, este supuesto suele ser razonable. En la sección 6-6 se presentará un método simple para verificar este supuesto, y se analizarán las acciones que deberán emprenderse en caso de que se viole.

6-2 EL DISEÑO 2^2

El primer diseño de la serie 2^k es el que sólo tiene dos factores, por ejemplo, A y B; cada uno se corre a dos niveles. A este diseño se le llama **diseño factorial 2^2**. Los niveles de los factores pueden denominarse arbitrariamente "bajo" y "alto". Como un ejemplo, considere la investigación del efecto de la concentración del reactivo y de la cantidad del catalizador sobre la conversión (rendimiento) de un proceso químico. Sea la concentración del reactivo el factor A, y sean 15 y 25 por ciento los dos niveles de interés. El catalizador es el factor B, con el nivel alto denotando el uso de 2 libras del catalizador y el nivel bajo denotando el uso de 1 libra. Se hacen tres réplicas del experimento, y los datos son los siguientes:

Factor		Combinación de tratamientos	Réplica			
A	B		I	II	III	Total
−	−	A bajo, B bajo	28	25	27	80
+	−	A alto, B bajo	36	32	32	100
−	+	A bajo, B alto	18	19	23	60
+	+	A alto, B alto	31	30	29	90

Las combinaciones de los tratamientos se ilustran gráficamente en la figura 6-1. Por convención, el efecto de un factor se denota con una letra mayúscula latina. Por lo tanto, "A" se refiere al efecto del factor A, "B" al efecto del factor B, y "AB" a la interacción AB. En el diseño 2^2, los niveles bajo y alto de A y B se denotan por "−" y "+", respectivamente, en los ejes A y B. Por lo tanto, − en el eje A representa el nivel bajo de la concentración (15%), mientras que + representa el nivel alto (25%), y − en el eje B representa el nivel bajo del catalizador, mientras que + denota el nivel alto.

Las cuatro combinaciones de tratamientos suelen representarse con letras minúsculas, como se muestra en la figura 6-1. Por la figura se observa que el nivel alto de cualquiera de los factores en una combinación de tratamientos se denota por la letra minúscula correspondiente y que el nivel bajo de un factor en una combinación de tratamientos se denota por la ausencia de la letra respectiva. Por lo tanto, a repre-

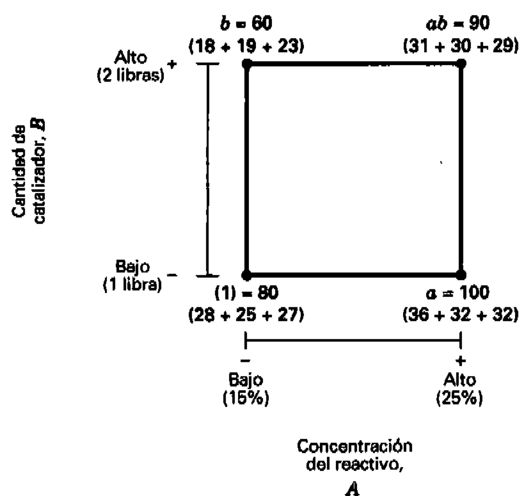


Figura 6-1 Combinaciones de los tratamientos en el diseño 2^2.

senta la combinación de tratamientos con A en el nivel alto y B en el nivel bajo, b representa A en el nivel bajo y B en el nivel alto, y ab representa ambos factores en el nivel alto. Por convención, se usa (1) para denotar que ambos factores están en el nivel bajo. Esta notación se utiliza en todas las series 2^k.

En un diseño factorial con dos niveles, el efecto promedio de un factor puede definirse como el cambio en la respuesta producido por un cambio en el nivel de ese factor promediado para los niveles del otro factor. Asimismo, los símbolos (1), a, b y ab representan ahora el *total* de las n réplicas hechas con la combinación de los tratamientos, como se ilustra en la figura 6-1. Ahora el efecto de A en el nivel bajo de B es $[a-(1)]/n$ y el efecto de A con el nivel alto de B es $[ab-b]/n$. Al promediarse estas dos cantidades se obtiene el **efecto principal** de A:

$$A=\frac{1}{2n}\{[ab-b]+[a-(1)]\}$$

$$=\frac{1}{2n}[ab+a-b-(1)] \tag{6-1}$$

El efecto principal promedio de B se encuentra a partir del efecto de B con el nivel bajo de A (es decir, $[b-(1)]/n$) y con el nivel alto de A (o sea, $[ab-a]/n$) como

$$B=\frac{1}{2n}\{[ab-a]+[b-(1)]\}$$

$$=\frac{1}{2n}[ab+b-a-(1)] \tag{6-2}$$

El **efecto de la interacción** AB se define como la diferencia promedio entre el efecto de A con el nivel alto de B y el efecto de A con el nivel bajo de B. Por lo tanto,

$$AB=\frac{1}{2n}\{[ab-b]-[a-(1)]\}$$

$$=\frac{1}{2n}[ab+(1)-a-b] \tag{6-3}$$

De manera alternativa, AB puede definirse como la diferencia promedio entre el efecto de B con el nivel alto de A y el efecto de B con el nivel bajo de A. Esto llevará también a la ecuación 6-3.

Las fórmulas de los efectos de A, B y AB pueden deducirse con otro método. El efecto de A puede encontrarse como la diferencia en la respuesta promedio de las dos combinaciones de tratamientos situadas a la derecha del cuadrado de la figura 6-1 (a este promedio se le llama $\bar{y}_{A^+}$, porque es la respuesta promedio con las combinaciones de tratamientos donde A está en el nivel alto) y las dos combinaciones de tratamientos situadas a la izquierda del cuadrado de la figura 6-1 (o $\bar{y}_{A^-}$). Es decir,

$$A=\bar{y}_{A^+}-\bar{y}_{A^-}$$

$$=\frac{ab+a}{2n}-\frac{b+(1)}{2n}$$

$$=\frac{1}{2n}[ab+a-b-(1)]$$

Se trata exactamente del mismo resultado que el de la ecuación 6-1. El efecto de B, ecuación 6-2, se encuentra como la diferencia entre el promedio de las dos combinaciones de tratamientos de la parte su-

perior del cuadrado ($\bar{y}_{B^+}$) y el promedio de las dos combinaciones de tratamientos de la parte inferior ($\bar{y}_{B^-}$), o

$$\begin{aligned} B &= \bar{y}_{B^+} - \bar{y}_{B^-} \\ &= \frac{ab+b}{2n} - \frac{a+(1)}{2n} \\ &= \frac{1}{2n}[ab+b-a-(1)] \end{aligned}$$

Por último, el efecto de la interacción AB es el promedio de las combinaciones de tratamientos de la diagonal de derecha a izquierda del cuadrado [ab y (1)] menos el promedio de las combinaciones de tratamientos de la diagonal de izquierda a derecha (a y b), o

$$\begin{aligned} AB &= \frac{ab+(1)}{2n} - \frac{a+b}{2n} \\ &= \frac{1}{2n}[ab+(1)-a-b] \end{aligned}$$

resultado que es idéntico a la ecuación 6-3.

Utilizando el experimento de la figura 6-1, los efectos promedio pueden estimarse como

$$\begin{aligned} A &= \frac{1}{2(3)}(90+100-60-80) = 8.33 \\ B &= \frac{1}{2(3)}(90+60-100-80) = -5.00 \\ AB &= \frac{1}{2(3)}(90+80-100-60) = 1.67 \end{aligned}$$

El efecto de A (concentración del reactivo) es positivo; esto sugiere que al incrementar A del nivel bajo (15%) al nivel alto (25%), el rendimiento se incrementará. El efecto de B (catalizador) es negativo; esto sugiere que al incrementar la cantidad del catalizador que se agrega al proceso se reducirá el rendimiento. El efecto de la interacción parece ser pequeño en comparación con los dos efectos principales.

En muchos experimentos que incluyen diseños 2^k, se examinará la **magnitud** y la **dirección** de los efectos de los factores a fin de determinar las variables que son de posible importancia. En la mayoría de los casos puede usarse el análisis de varianza para confirmar esta interpretación. Hay varios paquetes de software de estadística excelentes que son útiles para establecer y analizar diseños 2^k. Se cuenta también con métodos especiales que ahorran tiempo cuando los cálculos se hacen manualmente.

Considere las sumas de cuadrados de A, B y AB. Observe, por la ecuación 6-1, que se usó un **contraste** para estimar A, a saber

$$\text{Contraste}_A = ab+a-b-(1) \tag{6-4}$$

A este contraste suele llamársele el **efecto total** de A. A partir de las ecuaciones 6-2 y 6-3, se observa que también se usan contrastes para estimar B y AB. Además, estos tres contrastes son **ortogonales.** La suma de cuadrados de cualquier contraste puede calcularse con la ecuación 3-29, la cual establece que la suma de cuadrados del contraste es igual al cuadrado del contraste dividido por el número de observaciones en

cada total del contraste multiplicado por la suma de cuadrados de los coeficientes del contraste. Por consiguiente, se tienen

$$SS_A = \frac{[ab+a-b-(1)]^2}{4n} \tag{6-5}$$

$$SS_B = \frac{[ab+b-a-(1)]^2}{4n} \tag{6-6}$$

y

$$SS_{AB} = \frac{[ab+(1)-a-b]^2}{4n} \tag{6-7}$$

como las sumas de cuadrados de *A, B* y *AB*.

Al utilizar el experimento de la figura 6-1, las sumas de cuadrados de las ecuaciones 6-5, 6-6 y 6-7 pueden encontrarse como

$$SS_A = \frac{(50)^2}{4(3)} = 208.33$$
$$SS_B = \frac{(-30)^2}{4(3)} = 75.00 \tag{6-8}$$

y

$$SS_{AB} = \frac{(10)^2}{4(3)} = 8.33$$

La suma de cuadrados total se encuentra como de costumbre, es decir,

$$SS_T = \sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{n} y_{ijk}^2 - \frac{y_{...}^2}{4n} \tag{6-9}$$

En general, SS_T tiene $4n - 1$ grados de libertad. La suma de cuadrados del error, con $4(n-1)$ grados de libertad, suele calcularse por sustracción como

$$SS_E = SS_T - SS_A - SS_B - SS_{AB} \tag{6-10}$$

Para el experimento de la figura 6-1, se obtiene

$$SS_T = \sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{3} y_{ijk}^2 - \frac{y_{...}^2}{4(3)}$$
$$= 9398.00 - 9075.00 = 323.00$$

y

$$SS_E = SS_T - SS_A - SS_B - SS_{AB}$$
$$= 323.00 - 208.33 - 75.00 - 8.33$$
$$= 31.34$$

al utilizar SS_A, SS_B y SS_{AB} de la ecuación 6-8. En la tabla 6-1 se resume el análisis de varianza completo. Con base en los valores *P*, se concluye que los efectos principales son estadísticamente significativos y que no hay interacción entre estos factores. Esto confirma la interpretación de los datos que se hizo originalmente con base en las magnitudes de los efectos de los factores.

Con frecuencia resulta conveniente escribir las combinaciones de los tratamientos en el orden (1), *a*,

Tabla 6-1 Análisis de varianza del experimento de la figura 6-1

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0	Valor *P*
A	208.33	1	208.33	53.15	0.0001
B	75.00	1	75.00	19.13	0.0024
AB	8.33	1	8.33	2.13	0.1826
Error	31.34	8	3.92		
Total	323.00	11			

b, *ab*. Se hace referencia a esto como el **orden estándar** (u orden de Yates, por el Dr. Frank Yates). Al utilizar este orden estándar, se observa que los coeficientes de los contrastes usados para estimar los efectos son

Efectos	(1)	*a*	*b*	*ab*
A:	−1	+1	−1	+1
B:	−1	−1	+1	+1
AB:	+1	−1	−1	+1

Observe que los coeficientes de los contrastes para estimar el efecto de la interacción son sólo el producto de los coeficientes correspondientes de los dos efectos principales. El coeficiente de un contraste es siempre +1 o −1, y puede usarse una **tabla de signos positivos y negativos** como la tabla 6-2 para determinar el signo correcto para cada combinación de tratamientos. Los encabezados de las columnas de la tabla 6-2 son los efectos principales (*A* y *B*), la interacción *AB* e *I*, que representa el total o promedio del experimento completo. Observe que la columna que corresponde a *I* incluye únicamente signos positivos. Las etiquetas de los renglones son las combinaciones de los tratamientos. Para encontrar el contraste para estimar cualquier efecto, simplemente se multiplican los signos de la columna apropiada de la tabla por la combinación de tratamientos correspondiente y se hace la suma. Por ejemplo, para estimar *A*, el contraste es $-(1) + a - b + ab$, que concuerda con la ecuación 6-1.

El modelo de regresión

En un diseño factorial 2^k es sencillo expresar los resultados del experimento en términos de un **modelo de regresión.** Puesto que 2^k es tan sólo un diseño factorial, podría usarse un modelo de los efectos o de las medias, pero el enfoque del modelo de regresión es mucho más natural e intuitivo. Para el experimento del proceso químico de la figura 6-1, el modelo de regresión es

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$$

donde x_1 es una **variable codificada** que representa la concentración del reactivo y x_2 es una variable codificada que representa la cantidad del catalizador y las β son los coeficientes de regresión. La relación en-

Tabla 6-2 Signos algebraicos para calcular los efectos en el diseño 2^2

Combinación de tratamientos	Efecto factorial			
	I	*A*	*B*	*AB*
(1)	+	−	−	+
a	+	+	−	−
b	+	−	+	−
ab	+	+	+	+

tre las **variables naturales** —la concentración del reactivo y la cantidad de catalizador— y las variables codificadas es

$$x_1 = \frac{\text{Concentración} - (\text{Concentración}_{\text{baja}} + \text{Concentración}_{\text{alta}})/2}{(\text{Concentración}_{\text{alta}} - \text{Concentración}_{\text{baja}})/2}$$

y

$$x_2 = \frac{\text{Catalizador} - (\text{Catalizador}_{\text{bajo}} + \text{Catalizador}_{\text{alto}})/2}{(\text{Catalizador}_{\text{alto}} - \text{Catalizador}_{\text{bajo}})/2}$$

Cuando las variables naturales sólo tienen dos niveles, esta codificación producirá la familiar notación ± 1 para los niveles de las variables codificadas. Para ilustrar esto en el ejemplo tratado, observe que

$$x_1 = \frac{\text{Concentración} - (15+25)/2}{(25-15)/2} = \frac{\text{Concentración} - 20}{5}$$

Por lo tanto, si la concentración está en el nivel alto (Concentración = 25%), entonces $x_1 = +1$; si la concentración está en el nivel bajo (Concentración = 15%), entonces $x_1 = -1$. Además,

$$x_2 = \frac{\text{Catalizador} - (1+2)/2}{(2-1)/2} = \frac{\text{Catalizador} - 1.5}{0.5}$$

Por lo tanto, si el catalizador está en el nivel alto (Catalizador = 2 libras), entonces $x_2 = +1$; si el catalizador está en el nivel bajo (Catalizador = 1 libra), entonces $x_2 = -1$.

El modelo de regresión ajustado es

$$\hat{y} = 27.5 + \left(\frac{8.33}{2}\right)x_1 + \left(\frac{-5.00}{2}\right)x_2$$

donde la ordenada al origen es el gran promedio de las 12 observaciones, y los coeficientes de regresión $\hat{\beta}_1$ y $\hat{\beta}_2$ son la mitad de las estimaciones de los efectos de los factores correspondientes. La razón de que el coeficiente de regresión sea la mitad de la estimación del efecto es que un coeficiente de regresión mide el efecto de un cambio unitario en x sobre la media de y, y la estimación del efecto se basa en un cambio de dos unidades (de -1 a $+1$). Se demostrará más adelante que este método simple para estimar los coeficientes de regresión consiste en producir las estimaciones de **mínimos cuadrados** de los parámetros. Ver también el material suplementario de este capítulo.

Residuales y adecuación del modelo

El modelo de regresión puede usarse para obtener el valor predicho o ajustado de y en los cuatro puntos del diseño. Los residuales son las diferencias entre el valor observado y el valor ajustado de y. Por ejemplo, cuando la concentración del reactivo está en el nivel bajo ($x_1 = -1$) y el catalizador está en el nivel bajo ($x_2 = -1$), el rendimiento predicho es

$$\hat{y} = 27.5 + \left(\frac{8.33}{2}\right)(-1) + \left(\frac{-5.00}{2}\right)(-1) = 25.835$$

Hay tres observaciones en esta combinación de tratamientos, y los residuales son

$$e_1 = 28 - 25.835 = 2.165$$
$$e_2 = 25 - 25.835 = -0.835$$
$$e_3 = 27 - 25.835 = 1.165$$

Los valores predichos y los residuales restantes se calculan de manera similar. Para el nivel alto de la concentración del reactivo y el nivel bajo del catalizador,

$$\hat{y} = 27.5 + \left(\frac{8.33}{2}\right)(+1) + \left(\frac{-5.00}{2}\right)(-1)$$
$$= 34.165$$

y

$$e_4 = 36 - 34.165 = 1.835$$
$$e_5 = 32 - 34.165 = -2.165$$
$$e_6 = 32 - 34.165 = -2.165$$

Para el nivel bajo de la concentración del reactivo y el nivel alto del catalizador,

$$\hat{y} = 27.5 + \left(\frac{8.33}{2}\right)(-1) + \left(\frac{-5.00}{2}\right)(+1)$$
$$= 20.835$$

y

$$e_7 = 18 - 20.835 = -2.835$$
$$e_8 = 19 - 20.835 = -1.835$$
$$e_9 = 23 - 20.835 = 2.165$$

Por último, para el nivel alto de ambos factores,

$$\hat{y} = 27.5 + \left(\frac{8.33}{2}\right)(+1) + \left(\frac{-5.00}{2}\right)(+1)$$
$$= 29.165$$

y

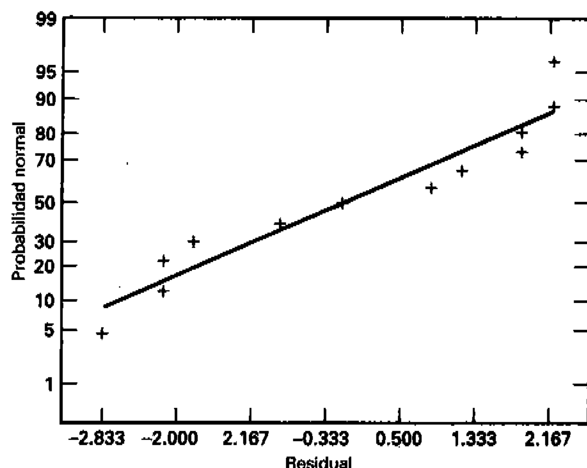
$$e_{10} = 31 - 29.165 = 1.835$$
$$e_{11} = 30 - 29.165 = 0.835$$
$$e_{12} = 29 - 29.165 = -0.165$$

En la figura 6-2 se presenta una gráfica de probabilidad normal de estos residuales y una gráfica de los residuales contra el rendimiento predicho. Estas gráficas parecen ser satisfactorias, por lo que no hay razón para sospechar problemas con la validez de las conclusiones.

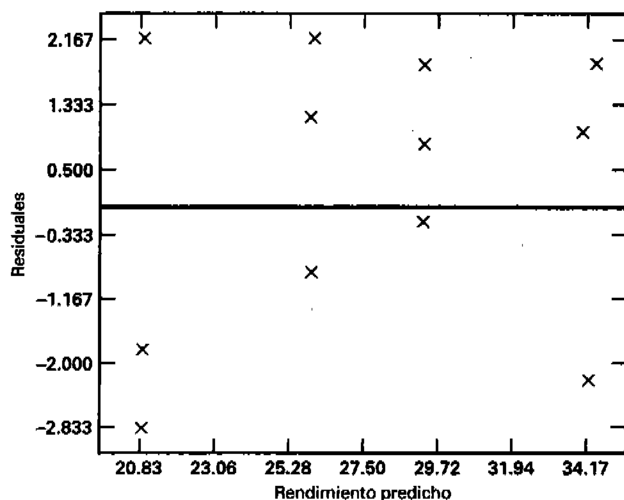
La superficie de respuesta

El modelo de regresión

$$\hat{y} = 27.5 + \left(\frac{8.33}{2}\right)x_1 + \left(\frac{-5.00}{2}\right)x_2$$



a) Gráfica de probabilidad normal



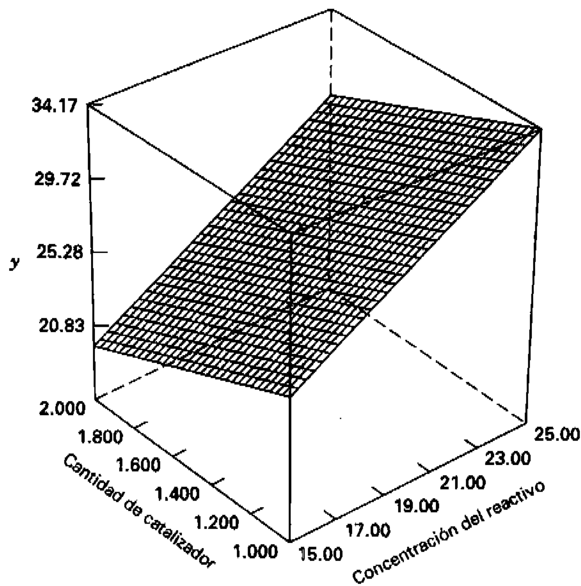
b) Residuales contra el rendimiento predicho

Figura 6-2 Gráficas de los residuales para el experimento del proceso químico.

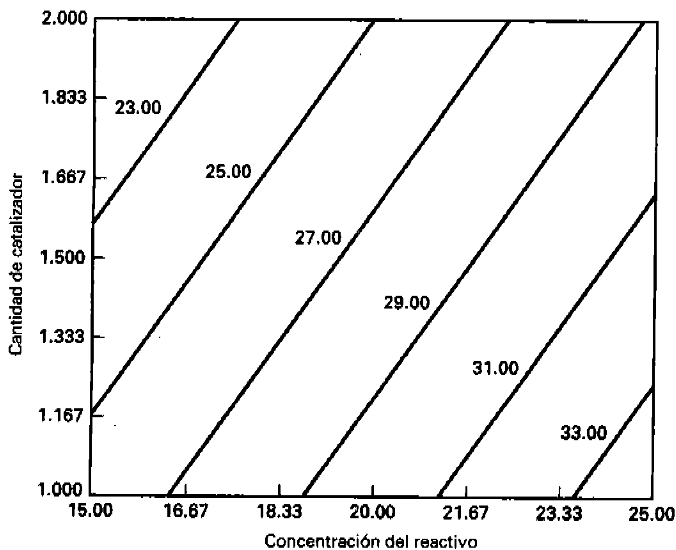
puede usarse para generar gráficas de superficie de respuesta. Si se desea construir estas gráficas en términos de los **niveles de los factores naturales,** entonces simplemente las relaciones entre las variables naturales y las codificadas que se dieron anteriormente se sustituyen en el modelo de regresión, de donde se obtiene

$$\hat{y} = 27.5 + \left(\frac{8.33}{2}\right)\left(\frac{\text{Concentración} - 20}{5}\right) + \left(\frac{-5.00}{2}\right)\left(\frac{\text{Catalizador} - 1.5}{0.5}\right)$$
$$= 18.33 + 0.8333\text{Concentración} - 5.00\text{Catalizador}$$

En la figura 6-3*a* se presenta la gráfica de superficie de respuesta tridimensional del rendimiento de este modelo, y la figura 6-3*b* es la gráfica de contorno. Puesto que el modelo es de **primer orden** (es decir, contiene únicamente los efectos principales), la superficie de respuesta ajustada es un plano. Al examinar



a) Superficie de respuesta



b) Gráfica de contorno

Figura 6-3 Gráfica de la superficie de respuesta y gráfica de contorno del experimento del proceso químico.

la gráfica de contorno se observa que el rendimiento aumenta cuando la concentración del reactivo se incrementa y la cantidad de catalizador disminuye. Frecuentemente se usa una superficie ajustada como ésta para encontrar la **dirección del mejoramiento potencial** de un proceso. Una manera formal de hacer esto, llamada **método del ascenso más pronunciado,** se presentará en el capítulo 11 cuando se estudien los métodos para realizar la exploración sistemática de las superficies de respuesta.

6-3 EL DISEÑO 2^3

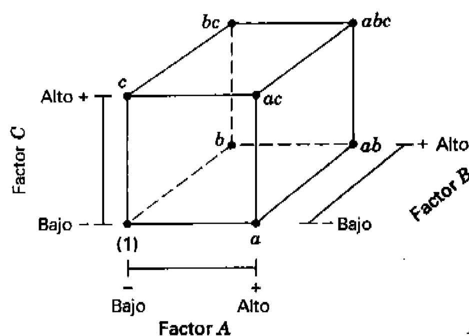
Suponga que tres factores, A, B y C, cada uno con dos niveles, son de interés. Al diseño se le llama **diseño factorial 2^3**, y en este caso la representación geométrica de las ocho combinaciones de tratamientos puede hacerse con un cubo, como se muestra en la figura 6-4*a*. Utilizando la notación "+" y "–" para representar los niveles alto y bajo de los factores, las ocho corridas del diseño 2^3 pueden enlistarse como en la figura 6-4*b*. Se le conoce en ocasiones como la **matriz del diseño.** Haciendo una ampliación de la notación de las etiquetas revisada en la sección 6-2, las combinaciones de los tratamientos en el orden estándar se escriben como (1), *a, b, ab, c, ac, bc* y *abc*. Recuerde que estos símbolos representan también el *total* de las n observaciones hechas con esa combinación de tratamientos particular.

Existen en realidad tres notaciones diferentes para las corridas del diseño 2^3 que son de uso general. La primera es la notación + y –, llamada con frecuencia **notación geométrica.** La segunda es el uso de las etiquetas en letras minúsculas para identificar las combinaciones de los tratamientos. La tercera y última notación utiliza 1 y 0 para denotar los niveles alto y bajo, respectivamente, de los factores, en lugar de + y –. Estas diferentes notaciones se ilustran enseguida para el diseño 2^3:

Corrida	A	B	C	Etiquetas	A	B	C
1	–	–	–	(1)	0	0	0
2	+	–	–	*a*	1	0	0
3	–	+	–	*b*	0	1	0
4	+	+	–	*ab*	1	1	0
5	–	–	+	*c*	0	0	1
6	+	–	+	*ac*	1	0	1
7	–	+	+	*bc*	0	1	1
8	+	+	+	*abc*	1	1	1

Hay siete grados de libertad entre las ocho combinaciones de tratamientos del diseño 2^3. Tres grados de libertad se asocian con los efectos principales de A, B y C. Cuatro grados de libertad se asocian con las interacciones; uno con cada una de las interacciones AB, AC y BC y uno con la interacción ABC.

Considere la estimación de los efectos principales. Primero, considere la estimación del efecto principal A. El efecto de A cuando B y C están en el nivel bajo es $[a - (1)]/n$. De manera similar, el efecto de A



a) Vista geométrica

	Factor		
Corrida	A	B	C
1	–	–	–
2	+	–	–
3	–	+	–
4	+	+	–
5	–	–	+
6	+	–	+
7	–	+	+
8	+	+	+

b) La matriz del diseño

Figura 6-4 El diseño factorial 2^3.

cuando B está en el nivel alto y C está en el nivel bajo es $[ab - b]/n$. El efecto de A cuando C está en el nivel alto y B está en el nivel bajo es $[ac - c]/n$. Por último, el efecto de A cuando tanto B como C están en el nivel alto es $[abc - bc]/n$. Por lo tanto, el efecto promedio de A es sólo el promedio de estos cuatro efectos, o

$$A = \frac{1}{4n}[a - (1) + ab - b + ac - c + abc - bc] \tag{6-11}$$

Esta ecuación también puede desarrollarse como un contraste entre las cuatro combinaciones de tratamientos de la cara derecha del cubo de la figura 6-5*a* (donde A está en el nivel alto) y las cuatro de la cara izquierda (donde A está en el nivel bajo). Es decir, el efecto de A es sólo el promedio de las cuatro corridas donde A está en el nivel alto ($\bar{y}_{A^+}$) menos el promedio de las cuatro corridas donde A está en el nivel bajo ($\bar{y}_{A^-}$), o

$$\begin{aligned} A &= \bar{y}_{A^+} - \bar{y}_{A^-} \\ &= \frac{a + ab + ac + abc}{4n} - \frac{(1) + b + c + bc}{4n} \end{aligned}$$

Esta ecuación puede reescribirse como

$$A = \frac{1}{4n}[a + ab + ac + abc - (1) - b - c - bc]$$

que es idéntica a la ecuación 6-11.

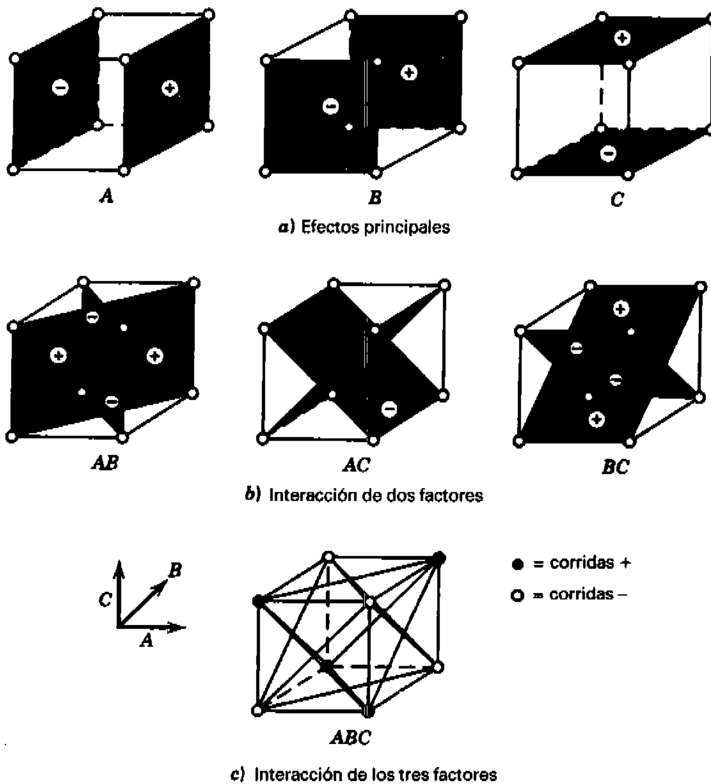


Figura 6-5 Representación geométrica de los contrastes que corresponden a los efectos principales y las interacciones del diseño 2^3.

De manera similar, el efecto de B es la diferencia en los promedios entre las cuatro combinaciones de tratamientos de la cara frontal del cubo y las cuatro de la cara posterior. Se obtiene así

$$\begin{aligned} B &= \bar{y}_{B^+} - \bar{y}_{B^-} \\ &= \frac{1}{4n}[b+ab+bc+abc-(1)-a-c-ac] \end{aligned} \tag{6-12}$$

El efecto de C es la diferencia en los promedios entre las cuatro combinaciones de tratamientos de la cara superior del cubo y las cuatro de la cara inferior, es decir,

$$\begin{aligned} C &= \bar{y}_{C^+} - \bar{y}_{C^-} \\ &= \frac{1}{4n}[c+ac+bc+abc-(1)-a-b-ab] \end{aligned} \tag{6-13}$$

Los efectos de la interacción de dos factores pueden calcularse con facilidad. Una medida de la interacción AB es la diferencia entre los efectos promedio de A con los dos niveles de B. Por convención, a la mitad de esta diferencia se le llama la interacción AB. Utilizando símbolos,

B	**Efecto promedio de *A***
Alto (+)	$\dfrac{[(abc-bc)+(ab-b)]}{2n}$
Bajo (–)	$\dfrac{\{(ac-c)+[a-(1)]\}}{2n}$
Diferencia	$\dfrac{[abc-bc+ab-b-ac+c-a+(1)]}{2n}$

Puesto que la interacción AB es la mitad de esta diferencia,

$$AB = \frac{[abc-bc+ab-b-ac+c-a+(1)]}{4n} \tag{6-14}$$

La ecuación 6-14 puede escribirse de la siguiente manera:

$$AB = \frac{abc+ab+c+(1)}{4n} - \frac{bc+b+ac+a}{4n}$$

En esta forma, resulta fácil ver que la interacción AB es la diferencia en los promedios entre las corridas de dos planos diagonales del cubo de la figura 6-5*b*. Utilizando un razonamiento lógico similar y con referencia a la figura 6-5*b*, las interacciones AC y BC son

$$AC = \frac{1}{4n}[(1)-a+b-ab-c+ac-bc+abc] \tag{6-15}$$

y

$$BC = \frac{1}{4n}[(1)+a-b-ab-c-ac+bc+abc] \tag{6-16}$$

La interacción ABC se define como la diferencia promedio entre la interacción AB para los dos diferentes niveles de C. Por lo tanto,

$$ABC = \frac{1}{4n}\{[abc - bc] - [ac - c] - [ab - b] + [a - (1)]\}$$
$$= \frac{1}{4n}[abc - bc - ac + c - ab + b + a - (1)] \quad (6\text{-}17)$$

Como antes, la interacción ABC puede considerarse como la diferencia de dos promedios. Si se aíslan las corridas de los dos promedios, éstas definen los vértices de los dos tetraedros que componen el cubo de la figura 6-5*c*.

En las ecuaciones 6-11 a 6-17, las cantidades entre corchetes son **contrastes** de las combinaciones de los tratamientos. Es posible desarrollar una tabla de signos positivos y negativos a partir de los contrastes, la cual se muestra en la tabla 6-3. Los signos de los efectos principales se determinan asociando un signo positivo con el nivel alto y un signo negativo con el nivel bajo. Una vez que se han establecido los signos de los efectos principales, los signos de las columnas restantes pueden obtenerse multiplicando las columnas precedentes apropiadas, renglón por renglón. Por ejemplo, los signos de la columna AB son el producto de los signos de la columna A y la columna B en cada renglón. El contraste de cualquier efecto puede obtenerse fácilmente con esta tabla.

La tabla 6-3 tiene varias propiedades interesantes: 1) Con excepción de la columna I, cada una de las columnas tienen el mismo número de signos positivos y negativos. 2) La suma de los productos de los signos de dos columnas cualesquiera es cero. 3) La columna I multiplicada por cualquiera de las columnas deja la columna sin cambio. Es decir, I es un **elemento identidad.** 4) El producto de dos columnas cualesquiera produce una columna de la tabla. Por ejemplo, $A \times B = AB$, y

$$AB \times B = AB^2 = A$$

Se observa que los exponentes de los productos se forman utilizando la aritmética **módulo 2.** (Es decir, el exponente sólo puede ser 0 o 1; si es mayor que 1, se reduce con múltiplos de 2 hasta que es 0 o 1.) Todas estas propiedades se derivan de la ortogonalidad de los contrastes usados para estimar los efectos.

Las sumas de cuadrados de los efectos se calculan con facilidad, ya que cada efecto tiene un contraste correspondiente con un solo grado de libertad. En el diseño 2^3 con n réplicas, la suma de cuadrados de cualquier efecto es

$$SS = \frac{(\text{Contraste})^2}{8n} \quad (6\text{-}18)$$

Tabla 6-3 Signos algebraicos para calcular los efectos del diseño 2^3

Combinación de tratamientos	Efecto factorial							
	I	A	B	AB	C	AC	BC	ABC
(1)	+	−	−	+	−	+	+	−
a	+	+	−	−	−	−	+	+
b	+	−	+	−	−	+	−	+
ab	+	+	+	+	−	−	−	−
c	+	−	−	+	+	−	−	+
ac	+	+	−	−	+	+	−	−
bc	+	−	+	−	+	−	+	−
abc	+	+	+	+	+	+	+	+

EJEMPLO 6-1 ··························

Recuerde el ejemplo 5-3, donde se presentó un estudio del efecto del porcentaje de carbonatación, la presión de operación y la velocidad de línea sobre la altura de llenado de una bebida carbonatada. Suponga que sólo se usan dos niveles de carbonatación, de tal modo que el experimento es un diseño factorial 2^3 con dos réplicas. Los datos (es decir, las desviaciones de la altura de llenado de especificación) se muestran en la tabla 6-4, y en la figura 6-6 se presenta la representación geométrica del diseño.

Al utilizar los totales bajo las combinaciones de los tratamientos que se muestran en la tabla 6-4, los efectos de los factores pueden estimarse de la siguiente manera:

$$\begin{aligned}
A &= \frac{1}{4n}[a-(1)+ab-b+ac-c+abc-bc] \\
&= \frac{1}{8}[1-(-4)+5-(-1)+3-(-1)+11-2] \\
&= \frac{1}{8}[24]=3.00 \\
B &= \frac{1}{4n}[b+ab+bc+abc-(1)-a-c-ac] \\
&= \frac{1}{8}[-1+5+2+11-(-4)-1-(-1)-3] \\
&= \frac{1}{8}[18]=2.25 \\
C &= \frac{1}{4n}[c+ac+bc+abc-(1)-a-b-ab] \\
&= \frac{1}{8}[-1+3+2+11-(-4)-1-(-1)-5] \\
&= \frac{1}{8}[14]=1.75 \\
AB &= \frac{1}{4n}[ab-a-b+(1)+abc-bc-ac+c] \\
&= \frac{1}{8}[5-1-(-1)+(-4)+11-2-3+(-1)] \\
&= \frac{1}{8}[6]=0.75
\end{aligned}$$

Tabla 6-4 El experimento de la altura de llenado, ejemplo 6-1

Corrida	Factores codificados *A*	*B*	*C*	Desviación de la altura de llenado: Réplica 1	Réplica 2	Niveles del factor	Bajo (−1)	Alto (+1)
1	−1	−1	−1	−3	−1	*A* (psi)	10	12
2	1	−1	−1	0	1	*B* (psi)	25	30
3	−1	1	−1	−1	0	*C* (b/min)	200	250
4	1	1	−1	2	3			
5	−1	−1	1	−1	0			
6	1	−1	1	2	1			
7	−1	1	1	1	1			
8	1	1	1	6	5			

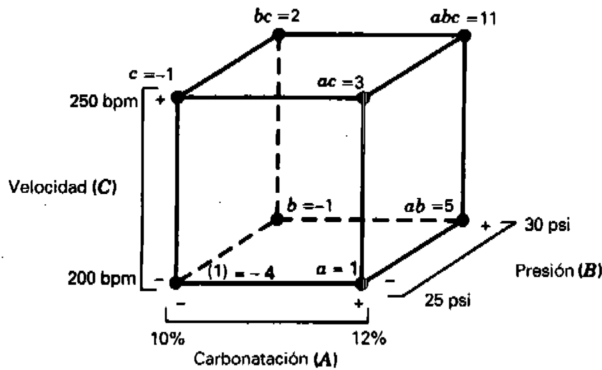


Figura 6-6 El diseño 2^3 para el experimento de la desviación de la altura de llenado del ejemplo 6-1.

$$
\begin{aligned}
AC &= \frac{1}{4n}[(1)-a+b-ab-c+ac-bc+abc] \\
&= \frac{1}{8}[-4-1+(-1)-5-(-1)+3-2+11] \\
&= \frac{1}{2}[2]=0.25 \\
BC &= \frac{1}{4n}[(1)+a-b-ab-c-ac+bc+abc] \\
&= \frac{1}{8}[-4+1-(-1)-5-(-1)-3+2+11] \\
&= \frac{1}{8}[4]=0.50
\end{aligned}
$$

y

$$
\begin{aligned}
ABC &= \frac{1}{4n}[abc-bc-ac+c-ab+b+a-(1)] \\
&= \frac{1}{8}[11-2-3+(-1)-5+(-1)+1-(-4)] \\
&= \frac{1}{8}[4]=0.50
\end{aligned}
$$

Los efectos más grandes son para la carbonatación (A = 3.00), la presión (B = 2.25), la velocidad (C = 1.75) y la interacción carbonatación-presión (AB = 0.75), si bien el efecto de la interacción no parece tener un impacto tan grande sobre la desviación de la altura de llenado como los efectos principales.

Las sumas de cuadrados se calculan con la ecuación 6-18 de la siguiente manera:

$$
\begin{aligned}
SS_A &= \frac{(24)^2}{16} = 36.00 \\
SS_B &= \frac{(18)^2}{16} = 20.25 \\
SS_C &= \frac{(14)^2}{16} = 12.25
\end{aligned}
$$

Tabla 6-5 Resumen de la estimación de los efectos del ejemplo 6-1

Factor	Estimación del efecto	Sumas de cuadrados	Contribución porcentual
A	3.00	36.00	46.1538
B	2.25	20.25	25.9615
C	1.75	12.25	15.7051
AB	0.75	2.25	2.88462
AC	0.25	0.25	0.320513
BC	0.50	1.00	1.28205
ABC	0.50	1.00	1.28205
Error puro		5.00	6.41026
Total		78.00	

$$SS_{AB} = \frac{(6)^2}{16} = 2.25$$

$$SS_{AC} = \frac{(2)^2}{16} = 0.25$$

$$SS_{BC} = \frac{(4)^2}{16} = 1.00$$

y

$$SS_{ABC} = \frac{(4)^2}{16} = 1.00$$

La suma de cuadrados total es $SS_T = 78.00$, y por sustracción, $SS_E = 5.00$. En la tabla 6-5 se resumen las estimaciones de los efectos y las sumas de cuadrados. La columna etiquetada "contribución porcentual" mide la contribución porcentual de cada uno de los términos del modelo a la suma de cuadrados total. La contribución porcentual es con frecuencia una guía aproximada pero efectiva de la importancia relativa de cada término del modelo. Observe que los efectos principales dominan en realidad este proceso, explicando más de 87% de la variabilidad total, mientras que la interacción AB explica menos de 3%.

El análisis de varianza de la tabla 6-6 puede usarse para confirmar la magnitud de estos efectos. Por la tabla 6-6 se observa que los efectos principales son altamente significativos (todos tienen valores P muy

Tabla 6-6 Análisis de varianza de los datos de la altura de llenado

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0	Valor P
Porcentaje de carbonatación (A)	36.00	1	36.00	57.60	<0.0001
Presión (B)	20.25	1	20.25	32.40	0.0005
Velocidad de línea (C)	12.25	1	12.25	19.60	0.0022
AB	2.25	1	2.25	3.60	0.0943
AC	0.25	1	0.25	0.40	0.5447
BC	1.00	1	1.00	1.60	0.2415
ABC	1.00	1	1.00	1.60	0.2415
Error	5.00	8	0.625		
Total	78.00	15			

pequeños). La interacción *AB* es significativa con un nivel aproximado de 10%; por lo tanto, existe una ligera interacción entre la carbonatación y la presión.

Quizá el lector quiera referirse al ejemplo 5-3 para la interpretación práctica de este experimento. Los responsables del proceso decidieron correrlo con presión baja y velocidad de línea alta, y reducir la variabilidad de la carbonatación controlando con mayor precisión la temperatura. Se consiguió así una reducción sustancial en la desviación de la altura de llenado del valor objetivo.

El modelo de regresión y la superficie de respuesta

El modelo de regresión para predecir la desviación de la altura de llenado es

$$\begin{aligned}\hat{y} &= \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \hat{\beta}_3 x_3 + \hat{\beta}_{12} x_1 x_2 \\ &= 1.00 + \left(\frac{3.00}{2}\right) x_1 + \left(\frac{2.25}{2}\right) x_2 + \left(\frac{1.75}{2}\right) x_3 + \left(\frac{0.75}{2}\right) x_1 x_2,\end{aligned}$$

donde las variables codificadas x_1, x_2 y x_3 representan a *A*, *B* y *C*, respectivamente. El término x_1x_2 es la interacción *AB*. Los residuales pueden obtenerse como la diferencia entre las desviaciones de la altura de llenado observada y la predicha. El análisis de estos residuales se deja como ejercicio para el lector.

En la figura 6-7 se muestra la superficie de respuesta y la gráfica de contorno para la desviación de la altura de llenado obtenida con el modelo de regresión, suponiendo que la velocidad de línea está en el nivel alto ($x_3 = 1$). Observe que como el modelo contiene la interacción, las líneas de contorno de la desviación de las alturas de llenado constantes son curvas (o la superficie de respuesta es un plano "torcido"). Es deseable operar este proceso de llenado de tal modo que la desviación del llenado esté tan cerca de cero como sea posible. La gráfica de contorno indica que si la velocidad de línea está en el nivel alto, entonces hay varias combinaciones de los niveles de la carbonatación y la presión que satisfarán este objetivo. Sin embargo, será necesario ejercer un control preciso de estas dos variables.

Solución por computadora

Hay muchos paquetes de software de estadística que establecerán y analizarán diseños factoriales con dos niveles. La salida de uno de estos programas de computadora, *Design-Expert*, se muestra en la tabla 6-7. En la parte superior de la tabla se presenta el análisis de varianza del modelo completo. El formato de esta presentación es un tanto diferente de los resultados dados en la tabla 6-6. Observe que el primer renglón del análisis de varianza es un resumen global del modelo completo (todos los efectos principales y las interacciones), y la suma de cuadrados del modelo es

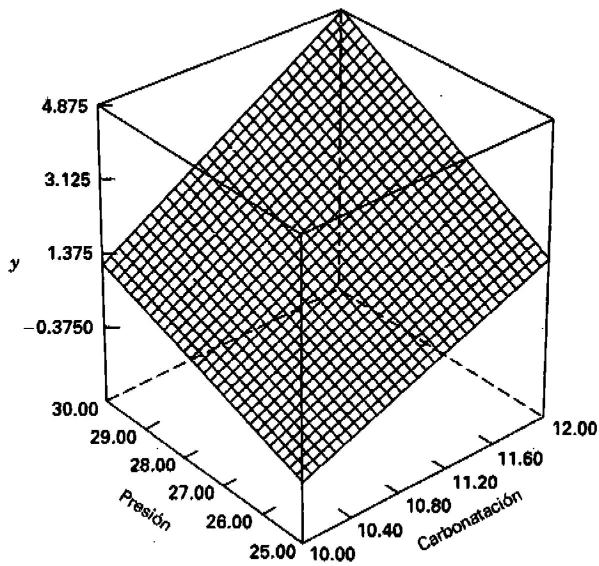
$$\begin{aligned}SS_{\text{Modelo}} &= SS_A + SS_B + SS_C + SS_{AB} + SS_{AC} + SS_{BC} + SS_{ABC} \\ &= 73.0\end{aligned}$$

Por lo tanto, el estadístico

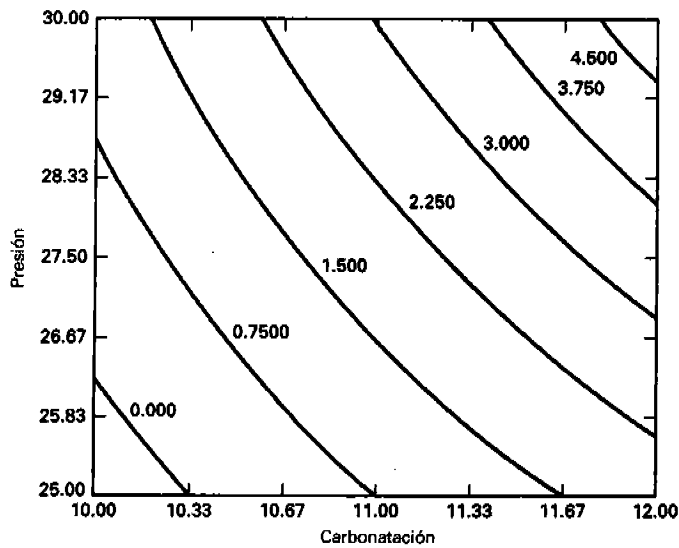
$$F_0 = \frac{MS_{\text{Modelo}}}{MS_E} = \frac{10.43}{0.63} = 16.69$$

está probando las hipótesis

$$\begin{aligned}&H_0: \beta_1 = \beta_2 = \beta_3 = \beta_{12} = \beta_{13} = \beta_{23} = \beta_{123} = 0 \\ &H_1: \text{al menos una } \beta \neq 0\end{aligned}$$



a) La superficie de respuesta



b) La gráfica de contorno

Figura 6-7 Superficie de respuesta y gráfica de contorno de la desviación de la altura de llenado, con la velocidad en el nivel alto (250 bpm), ejemplo 6-1.

Tabla 6-7 Salida de *Design-Expert* para el ejemplo 6-1

Response: Fill Deviation in Height
ANOVA for Selected Factorial Model
Analysis of variance table [Partial sum of squares]

Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	73.00	7	10.43	16.69	0.0003	significant
A	*36.00*	*1*	*36.00*	*57.60*	*<0.0001*	
B	*20.25*	*1*	*20.25*	*32.40*	*0.0005*	
C	*12.25*	*1*	*12.25*	*19.60*	*0.0022*	
AB	*2.25*	*1*	*2.25*	*3.60*	*0.0943*	
AC	*0.25*	*1*	*0.25*	*0.40*	*0.5447*	
BC	*1.00*	*1*	*1.00*	*1.60*	*0.2415*	
ABC	*1.00*	*1*	*1.00*	*1.60*	*0.2415*	
Residual	5.00	8	0.63			
Lack of Fit	*0.000*	*0*				
Pure Error	*5.00*	*8*	*0.63*			
Cor Total	78.00	15				

Std. Dev.	0.79	R-Squared	0.9359
Mean	1.00	Adj R-Squared	0.8798
C.V.	79.06	Pred R-Squared	0.7436
PRESS	20.00	Adeq Precision	13.416

Factor	Coefficient Estimate	DF	Standard Error	95% CI Low	95% CI High	VIF
Intercept	1.00	1	0.20	0.54	1.46	
A-Carbonation	1.50	1	0.20	1.04	1.96	1.00
B-Pressure	1.13	1	0.20	0.67	1.58	1.00
C-Speed	0.88	1	0.20	0.42	1.33	1.00
AB	0.38	1	0.20	−0.081	0.83	1.00
AC	0.13	1	0.20	−0.33	0.58	1.00
BC	0.25	1	0.20	−0.21	0.71	1.00
ABC	0.25	1	0.20	−0.21	0.71	1.00

Final Equation in Terms of Coded Factors:

Fill Deviation =
+1.00
+1.50 *A
+1.13 *B
+0.88 *C
+0.38 *A *B
+0.13 *A *C
+0.25 *B *C
+0.25 *A *B *C

Final Equation in Terms of Actual Factors:

Fill Deviation =
−225.50000
+21.00000 *Carbonation
+7.80000 *Pressure
+1.08000 *Speed
−0.75000 *Carbonation *Pressure
−0.10500 *Carbonation *Speed
−0.040000 *Pressure *Speed
+4.00000E-00: *Carbonation *Pressure *Speed

Reduced Model:
Response: Fill Deviation in Height
ANOVA for Selected Factorial Model
Analysis of variance table [Partial sum of squares]

Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	70.75	4	17.69	26.84	<0.0001	significant
A	*36.00*	*1*	*36.00*	*54.62*	*<0.0001*	
B	*20.25*	*1*	*20.25*	*30.72*	*0.0002*	
C	*12.25*	*1*	*12.25*	*18.59*	*0.0012*	
AB	*2.25*	*1*	*2.25*	*3.41*	*0.0917*	
Residual	7.25	11	0.66			
Lack of Fit	*2.25*	*3*	*0.75*	*1.20*	*0.3700*	*not significant*
Pure Error	*5.00*	*8*	*0.63*			
Cor Total	78.00	15				

Std. Dev.	0.81	R-Squared	0.9071
Mean	1.00	Adj R-Squared	0.8733
C.V.	81.18	Pred R-Squared	0.8033
PRESS	15.34	Adeq Precision	15.424

Factor	Coefficient Estimate	DF	Standard Error	95% CI Low	95% CI High	VIF
Intercept	1.00	1	0.20	0.55	1.45	
A-Carbonation	1.50	1	0.20	1.05	1.95	1.00
B-Pressure	1.13	1	0.20	0.68	1.57	1.00
C-Speed	0.88	1	0.20	0.43	1.32	1.00
AB	0.38	1	0.20	−0.072	0.82	1.00

Final Equation in Terms of Coded Factors:

Fill Deviation =
+1.00
+1.50 *A
+1.13 *B
+0.88 *C
+0.38 *A *B

Final Equation in Terms of Actual Factors:

Fill Deviation =
+9.62500
−2.62500 *Carbonation
−1.20000 *Pressure
+0.035000 *Speed
+0.15000 *Carbonation *Pressure

Diagnostics Case Statistics

Standard Order	Actual Value	Predicted Value	Residual	Leverage	Student Residual	Cook's Distance	Outlier t
1	−3.00	−2.13	−0.88	0.313	−1.300	0.154	−1.347
2	−1.00	−2.13	1.13	0.313	1.671	0.254	1.845
3	0.000	0.12	−0.12	0.313	−0.186	0.003	−0.177
4	1.00	0.12	0.88	0.313	1.300	0.154	1.347
5	−1.00	−0.63	−0.38	0.313	−0.557	0.028	−0.539
6	0.000	−0.63	0.63	0.313	0.928	0.078	0.922
7	2.00	3.13	−1.13	0.313	−1.671	0.254	−1.845
8	3.00	3.13	−0.13	0.313	−0.186	0.003	−0.177
9	−1.00	−0.37	−0.63	0.313	−0.928	0.078	−0.922
10	0.000	−0.37	0.37	0.313	0.557	0.028	0.539
11	2.00	1.88	0.13	0.313	0.186	0.003	0.177
12	1.00	1.88	−0.88	0.313	−1.300	0.154	−1.347
13	1.00	1.13	−0.13	0.313	−0.186	0.003	−0.177
14	1.00	1.13	−0.13	0.313	−0.186	0.003	−0.177
15	6.00	4.88	1.13	0.313	1.671	0.254	1.845
16	5.00	4.88	0.13	0.313	0.186	0.003	0.177

Puesto que F_0 es grande, se concluiría que al menos una de las variables tiene un efecto diferente de cero. Entonces se prueba la significación de cada efecto factorial individual utilizando el estadístico F. Estos resultados concuerdan con la tabla 6-6.

Abajo del análisis de varianza del modelo completo se presentan varios estadísticos R^2. La R^2 ordinaria es

$$R^2 = \frac{SS_{\text{Modelo}}}{SS_{\text{Total}}} = \frac{73.00}{78.00} = 0.9359$$

y mide la proporción de la variabilidad total explicada por el modelo. Un problema potencial con este estadístico es que siempre se incrementa cuando se agregan factores al modelo, incluso cuando estos factores no son significativos. El estadístico R^2 **ajustada,** definido como

$$R^2_{\text{Ajustada}} = 1 - \frac{SS_E / df_E}{SS_{\text{Total}} / df_{\text{Total}}} = 1 - \frac{5.00/8}{78.00/15} = 0.8798$$

es un estadístico que está ajustado para el "tamaño" del modelo; es decir, para el número de factores. La R^2 ajustada puede decrecer en realidad si se agregan términos no significativos al modelo. El estadístico PRESS es una medida de qué tan bien predecirá datos nuevos el modelo (PRESS es en realidad el acrónimo de *Prediction Error Sum of Squares* —suma de cuadrados del error de predicción—, y se calcula a partir de los errores de predicción obtenidos al predecir el punto i-ésimo de los datos con un modelo que incluye todas las observaciones, *excepto* la i-ésima). Un modelo con un valor pequeño de PRESS indica que es posible que el modelo sea un buen predictor. El estadístico "R^2 de predicción" se calcula como

$$R^2_{\text{Predicción}} = 1 - \frac{\text{PRESS}}{SS_{\text{Total}}} = 1 - \frac{20.00}{78.00} = 0.7436$$

Esto indica que se esperaría que el modelo completo explique cerca de 74% de la variabilidad de los datos nuevos.

La siguiente sección de la salida presenta el coeficiente de regresión de cada término del modelo y el **error estándar** (***se, standard error***) de cada coeficiente, definido como

$$se(\hat{\beta}) = \sqrt{V(\hat{\beta})} = \sqrt{\frac{MS_E}{n2^k}} = \sqrt{\frac{0.625}{2(8)}} = 0.20$$

Los intervalos de confianza de 95% para cada coeficiente de regresión se calculan a partir de

$$\hat{\beta} - t_{0.025,N-p}\,se(\hat{\beta}) \le \beta \le \hat{\beta} + t_{0.025,N-p}\,se(\hat{\beta})$$

donde los grados de libertad de t es el número de grados de libertad del error; es decir, N es el número total de corridas en el experimento (16), y p es el número de parámetros del modelo (8). También se presenta el modelo completo en términos de las variables codificadas y de las variables naturales.

En la última sección de la tabla 6-7 se ilustra la salida después de eliminar los términos de las interacciones no significativas. Este modelo reducido contiene ahora sólo los efectos principales A, B y C, y la interacción AB. La suma de cuadrados de los **residuales** o del **error** se compone ahora de un componente del **error puro** ("Pure Error") que surge de las réplicas de los ocho vértices del cubo, y un componente de **falta de ajuste** ("Lack of Fit"), compuesto por las sumas de cuadrados de las interacciones que se eliminaron del modelo (BC, AC y ABC). De nueva cuenta, la representación del modelo de regresión de los resultados experimentales se presenta en términos de las variables codificadas y las variables na-

turales. La proporción de la variabilidad total de la desviación de la altura del llenado que se explica por este modelo es

$$R^2 = \frac{SS_{\text{Modelo}}}{SS_{\text{Total}}} = \frac{70.75}{78.00} = 0.9071$$

que es menor que la R^2 del modelo completo. Observe, sin embargo, que la R^2 ajustada del modelo reducido apenas ha cambiado ligeramente respecto de la R^2 ajustada del modelo completo, y PRESS del modelo reducido es considerablemente menor, lo cual produce un valor más grande de $R^2_{\text{Predicción}}$ del modelo reducido. Evidentemente, la eliminación de los términos no significativos del modelo completo ha producido un modelo final que posiblemente funcionará con mayor eficiencia como predictor de datos nuevos. Observe que los intervalos de confianza para los coeficientes de regresión del modelo reducido son ligeramente más cortos que los intervalos de confianza correspondientes en el modelo completo.

En la última sección de la salida se presentan los residuales del modelo reducido. *Design-Expert* también construirá todas las gráficas de los residuales que se estudiaron anteriormente.

Otros métodos para evaluar la significación de los efectos

El análisis de varianza es una manera formal de determinar cuáles son los efectos de los factores que son diferentes de cero. Existen varios métodos más que son útiles. A continuación se indica cómo calcular el **error estándar de los efectos** y cómo usar los errores estándar para construir **intervalos de confianza** para los efectos. Otro método, que se ilustrará en la sección 6-5, utiliza **gráficas de probabilidad normal** para valorar la importancia de los efectos.

Es sencillo encontrar el error estándar de un efecto. Si se supone que hay n réplicas en cada una de las 2^k corridas del diseño, y si $y_{i1}, y_{i2}, \ldots, y_{in}$ son las observaciones de la corrida i-ésima, entonces

$$S_i^2 = \frac{1}{n-1}\sum_{j=1}^{n} (y_{ij} - \bar{y}_i)^2 \qquad i = 1, 2, \ldots, 2^k$$

es una estimación de la varianza de la corrida i-ésima. Las estimaciones de la varianza del diseño 2^k pueden combinarse para dar una estimación de la varianza global:

$$S^2 = \frac{1}{2^k(n-1)}\sum_{i=1}^{2^k}\sum_{j=1}^{n} (y_{ij} - \bar{y}_i)^2 \tag{6-19}$$

Ésta es también la estimación de la varianza dada por el cuadrado medio del error en el análisis de varianza. La *varianza* de la estimación de cada efecto es

$$\begin{aligned} V(\text{Efecto}) &= V\left(\frac{\text{Contraste}}{n2^{k-1}}\right) \\ &= \frac{1}{(n2^{k-1})^2} V(\text{Contraste}) \end{aligned}$$

Cada contraste es una combinación lineal de los 2^k totales de los tratamientos, y cada total consta de n observaciones. Por lo tanto,

$$V(\text{Contraste}) = n2^k \sigma^2$$

y la varianza de un efecto es

$$\begin{aligned} V(\text{Efecto}) &= \frac{1}{(n2^{k-1})^2} n2^k \sigma^2 \\ &= \frac{1}{n2^{k-2}} \sigma^2 \\ &= \frac{2\sigma}{\sqrt{n2^k}} \end{aligned}$$

El error estándar estimado se encontraría sacando la raíz cuadrada de esta última expresión y sustituyendo σ^2 con su estimación S^2:

$$se(\text{Efecto}) = \frac{2S}{\sqrt{n2^k}} \tag{6-20}$$

Observe que el error estándar de un efecto es el doble del error estándar de un coeficiente de regresión estimado en el modelo de regresión del diseño 2^k (ver la salida de computadora de *Design-Expert* del ejemplo 6-1).

Los intervalos de confianza de $100(1-\alpha)$ por ciento para los efectos se calculan a partir de Efecto $\pm$ $t_{\alpha/2,\,N-p}se(\text{Efecto})$, donde los grados de libertad de t son sólo los grados de libertad de los residuales o del error ($N-p$ = número total de corridas − número de parámetros del modelo).

Para ilustrar este método, considere el experimento de la desviación de la altura de llenado del ejemplo 6-1. El cuadrado medio del error es $MS_E = 0.625$. Por lo tanto, el error estándar de cada efecto es (utilizando $S^2 = MS_E$)

$$\begin{aligned} se(\text{Efecto}) &= \frac{2S}{\sqrt{n2^k}} \\ &= \frac{2\sqrt{0.625}}{\sqrt{2(2^3)}} \\ &= 0.40 \end{aligned}$$

Entonces, $t_{0.025,8} = 2.31$ y $t_{0.025,8}se(\text{Efecto}) = 2.31(0.40) = 0.92$, de donde los intervalos de confianza de 95% aproximados para los efectos de los factores son

$$\begin{aligned} A&: \ 3.00 \pm 0.92 \\ B&: \ 2.25 \pm 0.92 \\ C&: \ 1.75 \pm 0.92 \\ AB&: \ 0.75 \pm 0.92 \\ AC&: \ 0.25 \pm 0.92 \\ BC&: \ 0.50 \pm 0.92 \\ ABC&: \ 0.50 \pm 0.92 \end{aligned}$$

Este análisis indica que A, B y C son factores importantes, porque son las únicas estimaciones de los efectos de los factores para las que los intervalos de confianza de 95% aproximados no incluyen al cero.

Efectos de dispersión

El ingeniero de proceso que trabajó en el caso del llenado también se interesó en los **efectos de dispersión;** es decir, ¿alguno de los factores afecta la variabilidad de la desviación de la altura de llenado de una

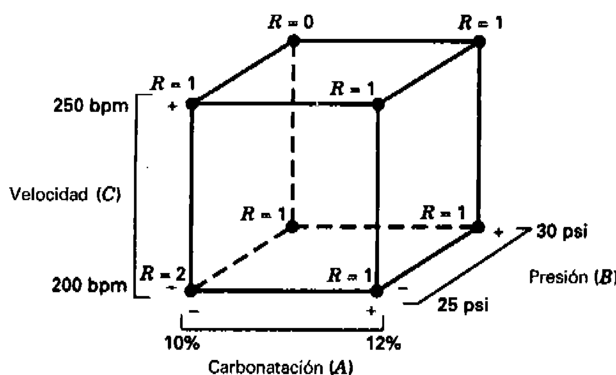


Figura 6-8 Rangos de la desviación de la altura de llenado del ejemplo 6-1.

corrida a otra? Una manera de responder esta pregunta es examinando el **rango** de las desviaciones de la altura de llenado para cada una de las ocho corridas del diseño 2^3. Estos rangos se grafican en el cubo de la figura 6-8. Observe que los rangos son aproximadamente iguales para las ocho corridas del diseño. Por consiguiente, no hay evidencia sólida que indique que alguna de las variables del proceso afecte directamente la variabilidad de la desviación de la altura de llenado en el proceso.

6-4 EL DISEÑO GENERAL 2^k

Los métodos de análisis que se han presentado hasta este punto pueden generalizarse para el caso de un **diseño factorial 2^k**, es decir, un diseño con k factores que tienen dos niveles cada uno. El modelo estadístico para un diseño 2^k incluiría k efectos principales, $\binom{k}{2}$ interacciones de dos factores, $\binom{k}{3}$ interacciones de tres factores, ..., y una interacción de k factores. Es decir, para un diseño 2^k el modelo completo contendría $2^k - 1$ efectos. También se usa aquí la notación introducida anteriormente para las combinaciones de los tratamientos. Por ejemplo, en un diseño 2^5, *abd* denota la combinación de tratamientos con los factores A, B y D en el nivel alto y los factores C y E en el nivel bajo. Las combinaciones de los tratamientos pueden escribirse en **orden estándar** introduciendo los factores uno a la vez y combinando sucesivamente cada nuevo factor con los que lo preceden. Por ejemplo, el orden estándar de un diseño 2^4 es (1), *a, b, ab, c, ac, bc, abc, d, ad, bd, abd, cd, acd, bcd* y *abcd.*

El enfoque general para el análisis estadístico del diseño 2^k se resume en la tabla 6-8. Como se señaló anteriormente, suele emplearse un paquete de software de computadora en este proceso de análisis.

A estas alturas, la secuencia de pasos de la tabla 6-8 debe resultar familiar. El primer paso es estimar los efectos de los factores y examinar sus signos y magnitudes. De este modo el experimentador obtiene in-

Tabla 6-8 Procedimiento de análisis para un diseño 2^k

1. Estimar los efectos de los factores
2. Formar el modelo inicial
3. Realizar las pruebas estadísticas
4. Refinar el modelo
5. Analizar los residuales
6. Interpretar los resultados

formación preliminar respecto de los factores y las interacciones que pueden ser importantes, y en qué direcciones deberán ajustarse estos factores para mejorar la respuesta. Para formar el modelo inicial del experimento, por lo general se elige el **modelo completo,** es decir, todos los efectos principales y las interacciones, siempre que se haya hecho una réplica de al menos uno de los puntos del diseño (en la sección siguiente se revisa una modificación de este paso). Después, en el paso 3 se usa el análisis de varianza para probar formalmente la significación de los efectos principales y las interacciones. En la tabla 6-9 se presenta la forma general de un análisis de varianza para un diseño factorial 2^k con n réplicas. El paso 4, refinar el modelo, suele consistir en la eliminación de las variables no significativas del modelo completo. El paso 5 es el análisis residual usual para verificar la adecuación del modelo y los supuestos. En ocasiones ocurrirá una refinación del modelo después del análisis residual, si se encuentra que el modelo es inadecuado o que hay violaciones serias de los supuestos. El último paso consiste generalmente en el análisis gráfico: gráficas de los efectos principales o las interacciones, o superficies de respuesta y gráficas de contorno.

Aun cuando los cálculos descritos se realizan por lo general con una computadora, en ocasiones es necesario calcular manualmente la estimación de un efecto o la suma de cuadrados de un efecto. Para estimar un efecto o calcular la suma de cuadrados de un efecto, primero debe determinarse el contraste asociado con ese efecto. Esto puede hacerse siempre utilizando una tabla de signos positivos y negativos, como la tabla 6-2 o 6-3. Sin embargo, para valores grandes de k esto resulta laborioso, y puede usarse un método alternativo. En general, el contraste del efecto $AB \cdots K$ se determina expandiendo el miembro derecho de

$$\text{Contraste}_{AB\cdots K} = (a \pm 1)(b \pm 1)\cdots(k \pm 1) \tag{6-21}$$

Tabla 6-9 Análisis de varianza de un diseño 2^k

Fuente de variación	Suma de cuadrados	Grados de libertad
k efectos principales		
A	SS_A	1
B	SS_B	1
⋮	⋮	⋮
K	SS_K	1
$\binom{k}{2}$ interacciones de dos factores		
AB	SS_{AB}	1
AC	SS_{AC}	1
⋮	⋮	⋮
JK	SS_{JK}	1
$\binom{k}{3}$ interacciones de tres factores		
ABC	SS_{ABC}	1
ABD	SS_{ABD}	1
⋮	⋮	⋮
IJK	SS_{IJK}	1
⋮	⋮	⋮
$\binom{k}{k} = 1$ interacción de k factores		
$ABC \cdots K$	$SS_{ABC\cdots K}$	1
Error	SS_E	$2^k(n-1)$
Total	SS_T	$n2^k - 1$

Para expandir la ecuación 6-21 se usa el álgebra ordinaria reemplazando "1" con (1) en la expresión final. El signo de cada grupo de paréntesis es negativo si el factor está incluido en el efecto y es positivo si el factor no está incluido.

Para ilustrar el uso de la ecuación 6-21, considere un diseño factorial 2^3. El contraste de AB sería

$$\begin{aligned}\text{Contraste}_{AB} &= (a-1)(b-1)(c+1)\\ &= abc+ab+c+(1)-ac-bc-a-b\end{aligned}$$

Como un ejemplo más, en un diseño 2^5, el contraste de $ABCD$ sería

$$\begin{aligned}\text{Contraste}_{ABCD} &= (a-1)(b-1)(c-1)(d-1)(e+1)\\ &= abcde+cde+bde+ade+bce\\ &\quad +ace+abe+e+abcd+cd+bd\\ &\quad +ad+bc+ac+ab+(1)-a-b-c\\ &\quad -abc-d-abd-acd-bcd-ae\\ &\quad -be-ce-abce-de-abde-acde-bcde\end{aligned}$$

Una vez que se han calculado los contrastes de los efectos, pueden estimarse los efectos y calcular las sumas de cuadrados de acuerdo con

$$AB\cdots K=\frac{2}{n2^k}(\text{Contraste}_{AB\cdots K}) \tag{6-22}$$

y

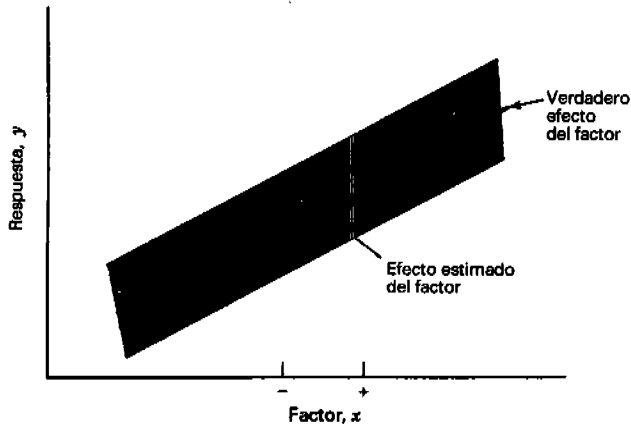
$$SS_{AB\cdots K}=\frac{1}{n2^k}(\text{Contraste}_{AB\cdots K})^2 \tag{6-23}$$

respectivamente, donde n denota el número de réplicas. Se cuenta también con un algoritmo tabular debido al Dr. Frank Yates que en ocasiones puede ser útil para el cálculo manual de las estimaciones de los efectos y las sumas de cuadrados. Referirse al material suplementario del texto de este capítulo.

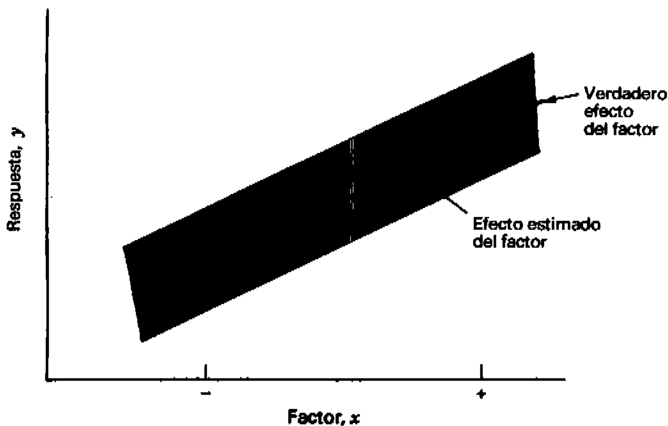
6-5 UNA SOLA RÉPLICA DEL DISEÑO 2^k

Incluso para un número moderado de factores, el número total de combinaciones de tratamientos en un diseño factorial 2^k es grande. Por ejemplo, un diseño 2^5 tiene 32 combinaciones de tratamientos, un diseño 2^6 tiene 64 combinaciones de tratamientos, etc. Debido a que por lo general los recursos son limitados, el número de réplicas que el experimentador puede emplear quizás esté restringido. Con frecuencia, los recursos disponibles permiten hacer únicamente una **sola réplica** del diseño, a menos que el experimentador esté dispuesto a omitir algunos de los factores originales.

Un riesgo obvio cuando se realiza un experimento que tiene una sola corrida para cada combinación de prueba es que el modelo puede ajustarse al ruido. Es decir, si la respuesta y es sumamente variable, pueden resultar conclusiones engañosas del experimento. La situación se ilustra en la figura 6-9*a*. En esta figura, la línea recta representa el verdadero efecto del factor. Sin embargo, debido a la variabilidad aleatoria presente en la variable de respuesta (representada por la franja sombreada), el experimentador obtiene en realidad las dos respuestas medidas representadas por los puntos negros. Por consiguiente, el efecto del factor estimado está cerca de cero y el experimentador ha llegado a una conclusión errónea respecto de este factor. Ahora bien, si hay menos variabilidad en la respuesta, la posibilidad de una conclusión errónea será más reducida. Otra forma de asegurarse de que se obtienen estimaciones confiables de los efectos es incrementando la distancia entre los niveles bajo (–) y alto (+) del factor, como se ilustra en



a) Distancia pequeña entre los niveles del factor



b) Separación agresiva de los niveles del factor

Figura 6-9 El impacto de la elección de los niveles del factor en un diseño no replicado.

la figura 6-9*b*. Observe que en esta figura la distancia incrementada entre los niveles bajo y alto del factor resulta en una estimación razonable del verdadero efecto del factor.

El uso de la estrategia de una sola réplica es común en los experimentos de exploración cuando hay un número relativamente grande de factores bajo consideración. Debido a que en estos casos nunca puede tenerse la certeza absoluta de que el error experimental es pequeño, una buena práctica en este tipo de experimentos es separar los niveles de los factores de manera agresiva. Quizás el lector encuentre útil releer las pautas generales para elegir los niveles de los factores del capítulo 1.

Una sola réplica de un diseño 2^k se denomina en ocasiones **diseño factorial no replicado.** Con una sola réplica, no se cuenta con ninguna estimación interna del error (o "error puro"). Una forma de abordar este análisis de un diseño factorial no replicado consiste en suponer que algunas interacciones de orden superior son insignificantes y combinar sus cuadrados medios para estimar el error. Esto es una apelación al **principio de efectos esparcidos;** es decir, la mayoría de los sistemas están dominados por algunos de los efectos principales y las interacciones de orden inferior, y la mayor parte de las interacciones de orden superior son insignificantes.

Cuando se analizan datos de diseños factoriales no replicados, ocasionalmente ocurren interacciones de orden superior reales. El uso de un cuadrado medio del error que se obtiene agrupando las interacciones de orden superior no es apropiado en estos casos. Un método de análisis atribuido a Daniel [35a] proporciona una forma simple de resolver este problema. Daniel sugiere examinar una **gráfica de probabilidad normal** de las estimaciones de los efectos. Los efectos que son insignificantes siguen una distribución normal, con media cero y varianza σ^2, y tenderán a localizarse sobre una línea recta en esta gráfica, mientras que los efectos significativos tendrán medias diferentes de cero y no se localizarán sobre la línea recta. Por lo tanto, el modelo preliminar se especificará de tal modo que contenga aquellos efectos que aparentemente son diferentes de cero, con base en la gráfica de probabilidad normal. Los efectos aparentemente insignificantes se combinan como una estimación del error.

EJEMPLO 6-2 ....

Una sola réplica del diseño 2^4

Un producto químico se fabrica en un envase presurizado. Se lleva a cabo un experimento factorial en la planta piloto para estudiar los factores que se piensa influyen en el índice de filtración de este producto. Los cuatro factores son la temperatura (A), la presión (B), la concentración del formaldehído (C) y la velocidad de agitación (D). Cada factor está presente con dos niveles. La matriz del diseño y los datos de la respuesta obtenidos de una sola réplica del experimento 2^4 se muestran en la tabla 6-10 y en la figura 6-10. Las 16 corridas se hacen de manera aleatoria. El ingeniero del proceso está interesado en maximizar el índice de filtración. Las condiciones actuales del proceso producen índices de filtración de alrededor de 75 gal/h. Asimismo, en el proceso actual la concentración de formaldehído, factor C, se usa en el nivel alto. Al ingeniero le gustaría reducir la concentración de formaldehído lo más posible, pero no ha podido hacerlo porque siempre produce índices de filtración más bajos.

El análisis de estos datos se iniciará construyendo una gráfica de probabilidad normal de las estimaciones de los efectos. La formación de signos positivos y negativos para las constantes de los contrastes

Tabla 6-10 Experimento del índice de filtración en la planta piloto

Número de corrida	Factor A	Factor B	Factor C	Factor D	Etiqueta de la corrida	Índice de filtración (gal/h)
1	−	−	−	−	(1)	45
2	+	−	−	−	*a*	71
3	−	+	−	−	*b*	48
4	+	+	−	−	*ab*	65
5	−	−	+	−	*c*	68
6	+	−	+	−	*ac*	60
7	−	+	+	−	*bc*	80
8	+	+	+	−	*abc*	65
9	−	−	−	+	*d*	43
10	+	−	−	+	*ad*	100
11	−	+	−	+	*bd*	45
12	+	+	−	+	*abd*	104
13	−	−	+	+	*cd*	75
14	+	−	+	+	*acd*	86
15	−	+	+	+	*bcd*	70
16	+	+	+	+	*abcd*	96

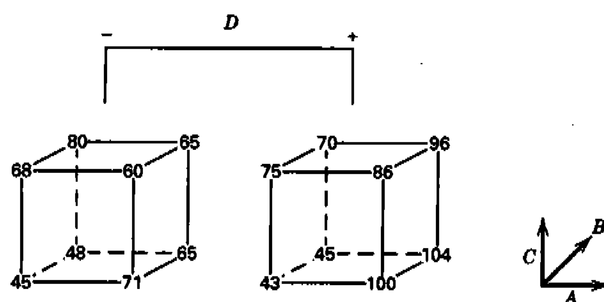


Figura 6-10 Datos del experimento del índice de filtración en la planta piloto para el ejemplo 6-2.

del diseño 2^4 se muestra en la tabla 6-11. A partir de estos contrastes pueden estimarse 15 efectos factoriales, y las sumas de cuadrados se presentan en la tabla 6-12.

En la figura 6-11 se muestra la gráfica de probabilidad normal de estos efectos. Todos los efectos que caen sobre la recta son insignificantes, mientras que los efectos grandes están apartados de ella. Los efectos importantes que surgen de este análisis son los efectos principales de A, C y D y las interacciones AC y AD.

Los efectos principales de A, C y D se grafican en la figura 6-12*a*. Los tres efectos son positivos, y si sólo se consideraran estos efectos principales, los tres factores se correrían en el nivel alto a fin de maximizar el índice de filtración. Sin embargo, siempre es necesario examinar cualquier interacción que sea importante. Recuerde que los efectos principales no tienen mucho significado cuando están presentes en interacciones significativas.

Las interacciones AC y AD se grafican en la figura 6-12*b*. Estas interacciones son la clave para resolver el problema. Observe, por la interacción AC, que el efecto de la temperatura es muy pequeño cuando la concentración está en el nivel alto y muy grande cuando la concentración está en el nivel bajo, obteniéndose los mejores resultados con la concentración baja y la temperatura alta. La interacción AD indica que la velocidad de agitación D tiene un efecto reducido con una temperatura baja, pero un efecto positivo grande con la temperatura alta. Por lo tanto, los mejores índices de filtración parecerían obtenerse cuando A y D están en el nivel alto y C está en el nivel bajo. Esto permitiría la reducción de la concentración de formaldehído a un nivel más bajo, otro de los objetivos del experimentador.

Proyección de un diseño

Es posible hacer otra interpretación de los efectos de la figura 6-11. Puesto que B (presión) no es significativa y todas las interacciones en las que interviene B son insignificantes, B puede descartarse del experimento, de tal modo que el diseño se convierte en un factorial 2^3 en A, C y D con dos réplicas. Esto es fácil de ver examinando únicamente las columnas A, C y D en la matriz del diseño que se muestra en la tabla 6-10 y observando que esas columnas forman dos réplicas de un diseño 2^3. En la tabla 6-13 se resume el análisis de varianza de los datos utilizando este supuesto de simplificación. Las conclusiones que se sacarían de este análisis se mantienen en esencia sin cambios respecto de las del ejemplo 6-2. Observe que al hacer la proyección de la réplica única del diseño 2^4 en un diseño 2^3 con dos réplicas, se tiene ahora tanto una estimación de la interacción ACD como una estimación del error basada en lo que en ocasiones se denomina **réplica oculta.**

Tabla 6-11 Constantes de los contrastes del diseño 2^4

	A	*B*	*AB*	*C*	*AC*	*BC*	*ABC*	*D*	*AD*	*BD*	*ABD*	*CD*	*ACD*	*BCD*	*ABCD*
(1)	−	−	+	−	+	+	−	−	+	+	−	+	−	−	+
a	+	−	−	−	−	+	+	−	−	+	+	+	+	−	−
b	−	+	−	−	+	−	+	−	+	−	+	+	−	+	−
ab	+	+	+	−	−	−	−	−	−	−	−	+	+	+	+
c	−	−	+	+	−	−	+	−	+	+	−	−	+	+	−
ac	+	−	−	+	+	−	−	−	−	+	+	−	−	+	+
bc	−	+	−	+	−	+	−	−	+	−	+	−	+	−	+
abc	+	+	+	+	+	+	+	−	−	−	−	−	−	−	−
d	−	−	+	−	+	+	−	+	−	−	+	−	+	+	−
ad	+	−	−	−	−	+	+	+	+	−	−	−	−	+	+
bd	−	+	−	−	+	−	+	+	−	+	−	−	+	−	+
abd	+	+	+	−	−	−	−	+	+	+	+	−	−	−	−
cd	−	−	+	+	−	−	+	+	−	−	+	+	−	−	+
acd	+	−	−	+	+	−	−	+	+	−	−	+	+	−	−
bcd	−	+	−	+	−	+	−	+	−	+	−	+	−	+	−
abcd	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

Tabla 6-12 Estimaciones de los efectos de los factores y sumas de cuadrados del diseño factorial 2^4 del ejemplo 6-2

Término del modelo	Estimación del efecto	Suma de cuadrados	Contribución porcentual
A	21.625	1870.56	32.6397
B	3.125	39.0625	0.681608
C	9.875	390.062	6.80626
D	14.625	855.563	14.9288
AB	0.125	0.0625	0.00109057
AC	−18.125	1314.06	22.9293
AD	16.625	1105.56	19.2911
BC	2.375	22.5625	0.393696
BD	−0.375	0.5625	0.00981515
CD	−1.125	5.0625	0.0883363
ABC	1.875	14.0625	0.245379
ABD	4.125	68.0625	1.18763
ACD	−1.625	10.5625	0.184307
BCD	−2.625	27.5625	0.480942
$ABCD$	1.375	7.5625	0.131959

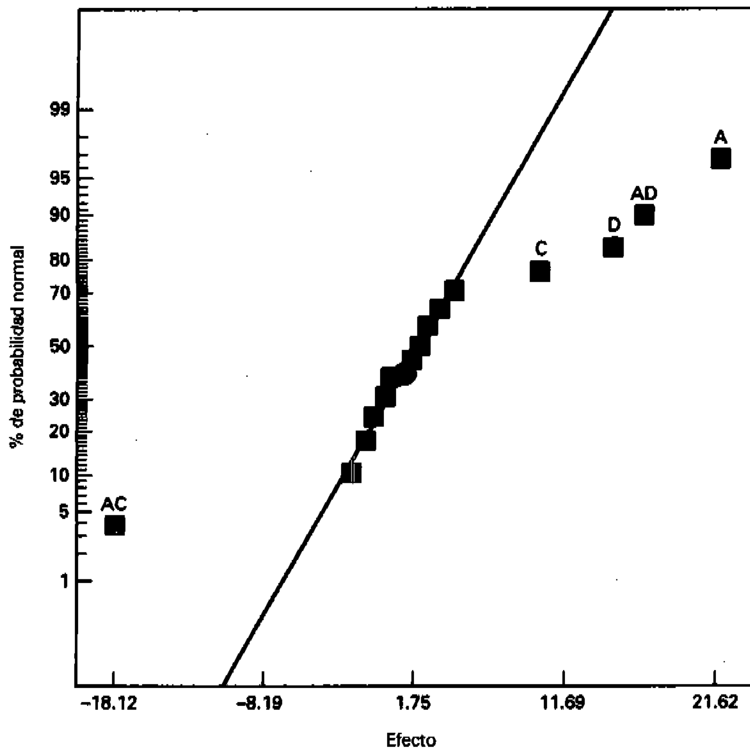
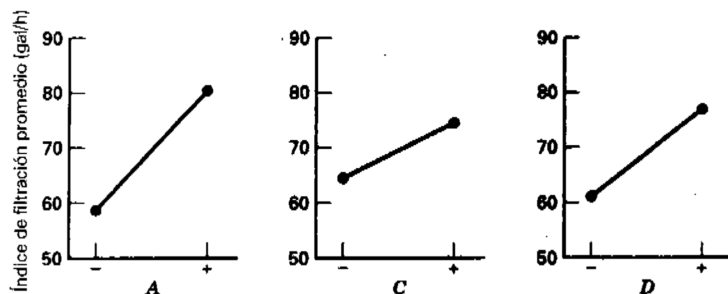
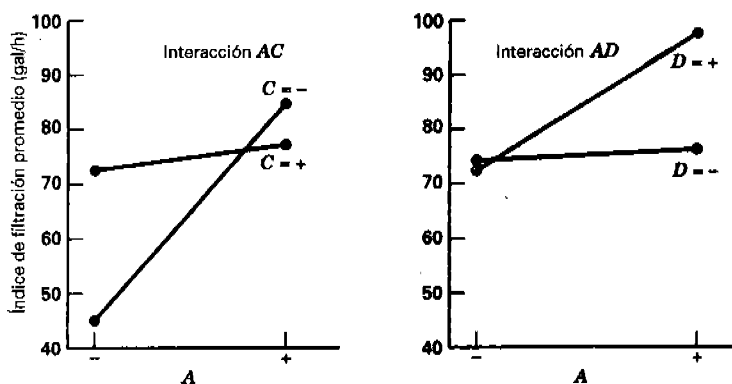


Figura 6-11 Gráfica de probabilidad normal de los efectos para el diseño factorial 2^4 del ejemplo 6-2.



a) Gráficas de los efectos principales



b) Gráficas de las interacciones

Figura 6-12 Gráficas de los efectos principales y las interacciones para el ejemplo 6-2.

Tabla 6-13 Análisis de varianza del experimento del índice filtración en la planta piloto en A, C y D

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0	Valor *P*
A	1870.56	1	1870.56	83.36	<0.0001
C	390.06	1	390.06	17.38	<0.0001
D	855.56	1	855.56	38.13	<0.0001
AC	1314.06	1	1314.06	58.56	<0.0001
AD	1105.56	1	1105.56	49.27	<0.0001
CD	5.06	1	5.06	<1	
ACD	10.56	1	10.56	<1	
Error	179.52	8	22.44		
Total	5730.94	15			

El concepto de proyectar un diseño factorial no replicado en un diseño factorial con réplicas en menos factores es muy útil. En general, si se tiene una sola réplica del diseño 2^k, y si h ($h < k$) factores son insignificantes y pueden descartarse, entonces los datos originales corresponden a un diseño factorial completo con dos niveles en los $k - h$ factores restantes con 2^h réplicas.

Verificación de diagnóstico

Deberán aplicarse las verificaciones de diagnóstico usuales a los residuales de un diseño 2^k. El análisis realizado indica que los únicos efectos significativos son $A = 21.625$, $C = 9.875$, $D = 14.625$, $AC = -18.125$ y $AD = 16.625$. Si esto es correcto, los índices de filtración estimados están dados por

$$\hat{y} = 70.06 + \left(\frac{21.625}{2}\right)x_1 + \left(\frac{9.875}{2}\right)x_3 + \left(\frac{14.625}{2}\right)x_4 - \left(\frac{18.125}{2}\right)x_1x_3 + \left(\frac{16.625}{2}\right)x_1x_4$$

donde 70.06 es la respuesta promedio y las variables codificadas x_1, x_3, x_4 asumen valores entre -1 y $+1$. El índice de filtración predicho para la corrida (1) es

$$\begin{aligned}\hat{y} &= 70.06 + \left(\frac{21.625}{2}\right)(-1) + \left(\frac{9.875}{2}\right)(-1) + \left(\frac{14.625}{2}\right)(-1) \\ &\quad - \left(\frac{18.125}{2}\right)(-1)(-1) + \left(\frac{16.625}{2}\right)(-1)(-1) \\ &= 46.22\end{aligned}$$

Puesto que el valor observado es 45, el residual es $e = y - \hat{y} = 45 - 46.22 = -1.22$. A continuación se presentan los valores de y, $\hat{y}$ y $e = y - \hat{y}$ para las 16 observaciones.

	y	$\hat{y}$	$e = y - \hat{y}$
(1)	45	46.22	−1.22
a	71	69.39	1.61
b	48	46.22	1.78
ab	65	69.39	−4.39
c	68	74.23	−6.23
ac	60	61.14	−1.14
bc	80	74.23	5.77
abc	65	61.14	3.86
d	43	44.22	−1.22
ad	100	100.65	−0.65
bd	45	44.22	0.78
abd	104	100.65	3.35
cd	75	72.23	2.77
acd	86	92.40	−6.40
bcd	70	72.23	−2.23
abcd	96	92.40	3.60

En la figura 6-13 se muestra la gráfica de probabilidad normal de los residuales. Los puntos de esta gráfica se localizan razonablemente próximos a una línea recta, brindando apoyo a la conclusión de que A, C, D, AC y AD son los únicos efectos significativos y que se satisfacen los supuestos fundamentales del análisis.

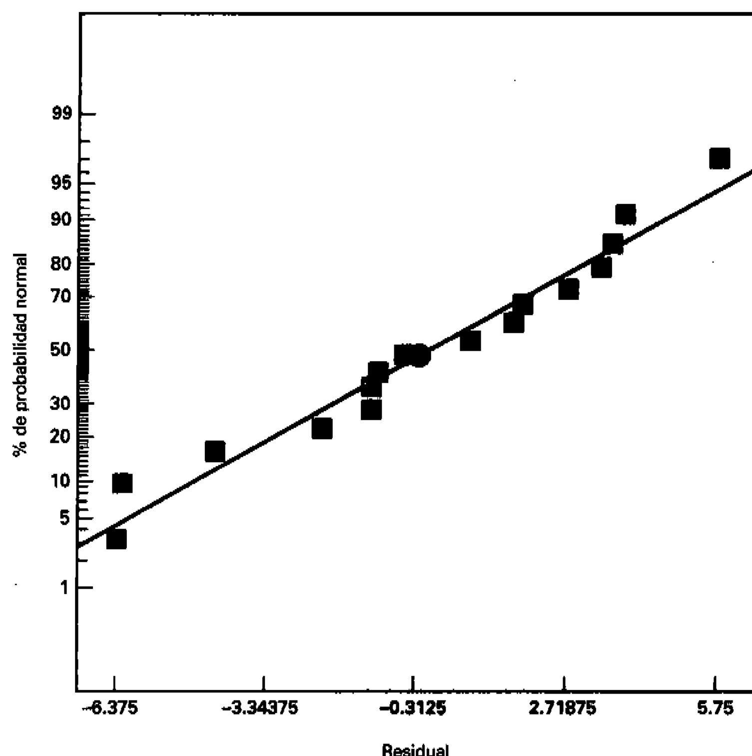


Figura 6-13 Gráfica de probabilidad normal de los residuales del ejemplo 6-2.

La superficie de respuesta

Las gráficas de las interacciones de la figura 6-12 se utilizaron para ofrecer una interpretación práctica de los resultados de este experimento. En ocasiones es útil emplear la superficie de respuesta para este fin. La superficie de respuesta se genera por el modelo de regresión

$$\hat{y} = 70.06 + \left(\frac{21.625}{2}\right)x_1 + \left(\frac{9.875}{2}\right)x_3 + \left(\frac{14.625}{2}\right)x_4 - \left(\frac{18.125}{2}\right)x_1x_3 + \left(\frac{16.625}{2}\right)x_1x_4$$

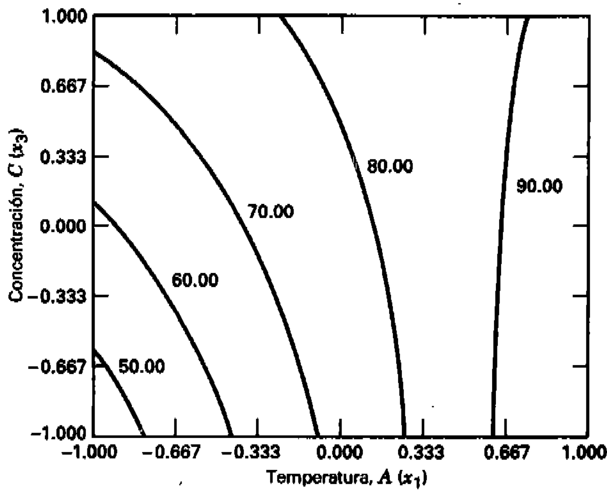
En la figura 6-14*a* se muestra la gráfica de contorno de la superficie de respuesta cuando la velocidad de agitación está en el nivel alto (es decir, $x_4 = 1$). Los contornos se generan a partir del modelo anterior con $x_4 = 1$, o

$$\hat{y} = 77.3725 + \left(\frac{38.25}{2}\right)x_1 + \left(\frac{9.875}{2}\right)x_3 - \left(\frac{18.125}{2}\right)x_1x_3$$

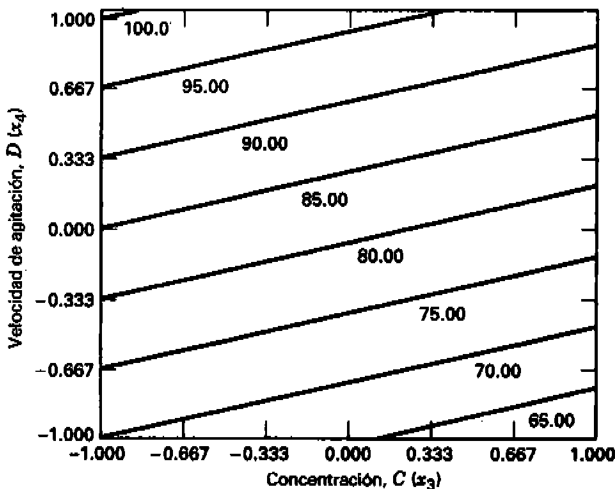
Observe que los contornos son líneas curvas porque el modelo contiene un término de interacción.

La figura 6-14*b* es la gráfica de contorno de la superficie de respuesta cuando la temperatura está en el nivel alto (es decir, $x_1 = 1$). Cuando se hace $x_1 = 1$ en el modelo de regresión se obtiene

$$\hat{y} = 80.8725 - \left(\frac{8.25}{2}\right)x_3 + \left(\frac{31.25}{2}\right)x_4$$



a) Gráfica de contorno con la velocidad de agitación (*D*), $x_4 = 1$



b) Gráfica de contorno con la temperatura (*A*), $x_1 = 1$

Figura 6-14 Gráficas de contorno del índice de filtración, ejemplo 6-2.

Estos contornos son rectas paralelas porque el modelo contiene únicamente los efectos principales de los factores C (x_3) y D (x_4).

Ambas gráficas de contorno indican que si se quiere maximizar el índice de filtración, las variables A (x_1) y D (x_4) deberán estar en el nivel alto y que el proceso es relativamente robusto para la concentración C. Se obtuvieron conclusiones similares a partir de las gráficas de las interacciones.

La mitad de gráfica normal de los efectos

Una alternativa para la gráfica de probabilidad normal de los efectos de los factores es la **mitad de gráfica normal.** Es una gráfica del valor absoluto de las estimaciones de los efectos contra sus probabilidades normales acumuladas. En la figura 6-15 se muestra la mitad de gráfica normal de los efectos para el ejemplo 6-2. La línea recta de la mitad de gráfica normal siempre pasa por el origen y deberá pasar también cerca del valor de los datos del percentil cincuenta. Muchos analistas sienten que es más fácil interpretar la mitad de

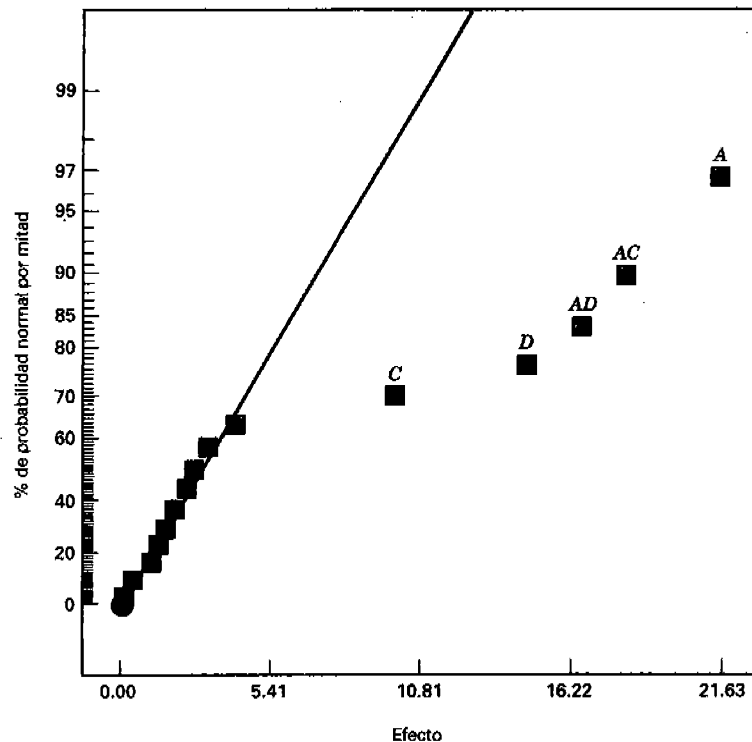


Figura 6-15 Mitad de gráfica normal de los efectos de los factores del ejemplo 6-2.

gráfica normal, en particular si sólo se cuenta con pocas estimaciones de los efectos, como cuando el experimentador ha usado un diseño de ocho corridas. Algunos paquetes de software construirán ambas gráficas.

Otros métodos para analizar diseños factoriales no replicados

El procedimiento de análisis estándar para un diseño factorial de dos factores no replicado es la gráfica normal (o mitad de gráfica normal) de los efectos estimados de los factores. Sin embargo, los diseños no replicados son tan usados en la práctica que se han propuesto muchos procedimientos formales de análisis para resolver la subjetividad de la gráfica de probabilidad normal. Hamada y Balakrishnan [52] compararon algunos de estos métodos. Encontraron que el método propuesto por Lenth [70] tiene una potencia adecuada para detectar efectos significativos. También es fácil de implementar y, como resultado, está empezando a aparecer en algunos paquetes de software para analizar datos de diseños factoriales no replicados. Se ofrece una breve descripción del método de Lenth.

Suponga que se tienen m contrastes de interés, por ejemplo $c_1, c_2, \ldots, c_m$. Si el diseño es un factorial 2^k no replicado, estos contrastes corresponden a las $m = 2^k - 1$ estimaciones de los efectos de los factores. La base del método de Lenth es estimar la varianza de un contraste a partir de las estimaciones más pequeñas (en valor absoluto) de los contrastes. Sean

$$s_0 = 1.5 \times \text{mediana} \ (|c_j|)$$

y

$$PSE = 1.5 \times \text{mediana} \ (|c_j| : |c_j| \le 2.5 s_0)$$

PSE denota el "pseudo error estándar", y Lenth demuestra que es un estimador razonable de la varianza del contraste cuando no hay muchos efectos activos (significativos). El *PSE* se usa para juzgar la significación de los contrastes. Un contraste individual puede compararse con el **margen de error** (***ME, margin of error***)

$$ME = t_{0.025,d} \times PSE$$

donde los grados de libertad se definen como $d = m/3$. Para hacer inferencias sobre un grupo de contrastes, Lenth sugiere usar el **margen de error simultáneo** (***SME, simultaneous margin of error***)

$$SME = t_{\gamma,d} \times PSE$$

donde el punto porcentual de la distribución t que se usa es $\gamma = 1 - (1 + 0.95^{1/m})/2$.

Para ilustrar el método de Lenth, considere el experimento 2^4 del ejemplo 6-2. Los cálculos dan como resultado $s_0 = 1.5 \times |-2.625| = 3.9375$ y $2.5 \times 3.9375 = 9.84375$, de donde

$$PSE = 1.5 \times |1.75| = 2.625$$
$$ME = 2.571 \times 2.625 = 6.75$$
$$SME = 5.219 \times 2.625 = 13.70$$

Considere ahora las estimaciones de los efectos de la tabla 6-12. El criterio *SME* indicaría que los cuatro efectos más grandes (en magnitud) son significativos, ya que las estimaciones de sus efectos exceden *SME*. El efecto principal de *C* es significativo de acuerdo con el criterio *ME*, pero no con respecto al *SME*. Sin embargo, puesto que es evidente que la interacción *AC* es importante, probablemente *C* se incluiría en la lista de efectos significativos. Observe que en este ejemplo el método de Lenth produjo la misma respuesta que la obtenida anteriormente con el examen de la gráfica de probabilidad normal de los efectos.

Varios autores (ver Hamada y Balakrishnan [52], Loughin [73], Loughin y Noble [74] y Larntz y Whitcomb [69]) han hecho notar que el método de Lenth falla para controlar los índices del error tipo I, y que pueden usarse métodos de simulación para calibrar su procedimiento. Larntz y Whitcomb [69] sugieren reemplazar los multiplicadores *ME* y *SME* con **multiplicadores ajustados** de la siguiente manera:

Número de contrastes	7	15	31
ME original	3.764	2.571	2.218
ME ajustado	2.295	2.140	2.082
SME original	9.008	5.219	4.218
SME ajustado	4.891	4.163	4.030

Estos resultados coinciden en gran medida con los de Ye y Hamada [114].

En general, el método de Lenth es un procedimiento ingenioso y útil. Sin embargo, recomendamos utilizarlo como **complemento** de la gráfica de probabilidad normal usual de los efectos, no como su sustituto.

Bisgaard [10] ha proporcionado una sutil técnica gráfica, llamada **carta de inferencia condicional,** como ayuda para interpretar la gráfica de probabilidad normal. La finalidad de esta gráfica es ayudar al experimentador a juzgar los efectos significativos. Esto sería relativamente sencillo si se conociera la desviación estándar σ, o si pudiera estimarse a partir de los datos. En diseños no replicados, no se cuenta con ninguna estimación interna de σ, por lo que la carta de inferencia condicional está diseñada para ayudar al experimentador a evaluar la magnitud de los efectos para un *rango* de valores de la desviación estándar.

Bisgaard fundamenta la gráfica en el resultado de que el error estándar de un efecto, en un diseño de dos niveles con N corridas (para un diseño factorial no replicado, $N = 2^k$), es

$$\frac{2\sigma}{\sqrt{N}}$$

donde σ es la desviación estándar de una observación individual. Entonces ± 2 veces el error estándar de un efecto es

$$\pm\frac{4\sigma}{\sqrt{N}}$$

Una vez que se estiman los efectos, se hace una gráfica como la que se muestra en la figura 6-16, con las estimaciones de los efectos graficadas en el eje vertical, o eje y. En esta figura se han usado las estimaciones de los efectos del ejemplo 6-2. El eje horizontal, o x, de la figura 6-16 es la escala de la desviación estándar (σ). Las dos rectas están en

$$y = +\frac{4\sigma}{\sqrt{N}} \quad \text{y} \quad y = -\frac{4\sigma}{\sqrt{N}}$$

En el ejemplo tratado aquí, $N = 16$, por lo que las rectas están en $y = +\sigma$ y $y = -\sigma$. Por lo tanto, para cualquier valor dado de la desviación estándar σ, la distancia entre estas dos rectas puede leerse como un intervalo de confianza de 95% aproximado para los efectos insignificantes.

En la figura 6-16 se observa que si el experimentador piensa que la desviación estándar está entre 4 y 8, entonces los factores A, C, D y las interacciones AC y AD son significativos. Si el experimentador piensa que la desviación estándar tiene un valor de hasta 10, el factor C quizá no sea significativo. Es decir, para

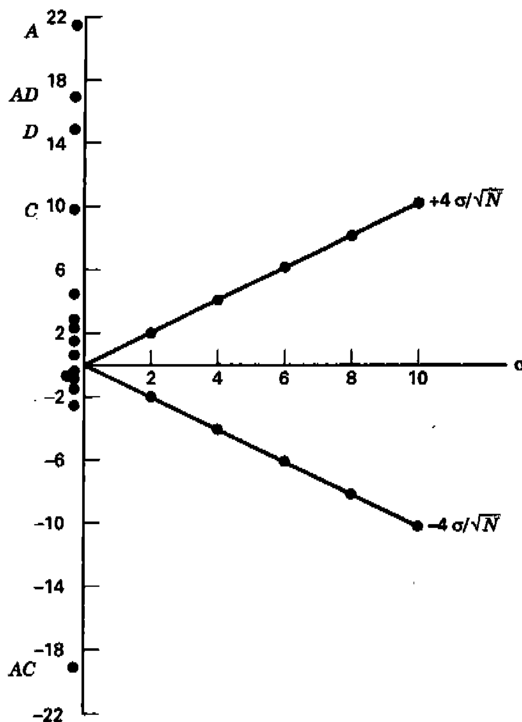


Figura 6-16 Carta de inferencia condicional para el ejemplo 6-2.

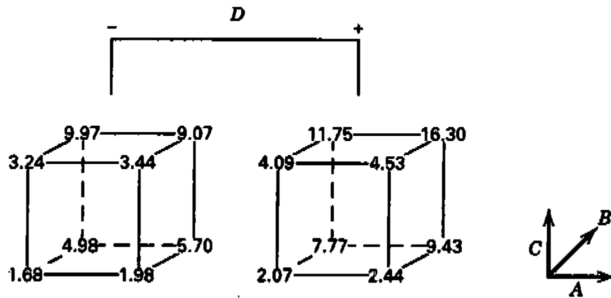


Figura 6-17 Datos del experimento de perforación del ejemplo 6-3.

cualquier supuesto dado acerca de la magnitud de σ, el experimentador puede construir una "cinta de medir" para juzgar la significación aproximada de los efectos. La carta también puede usarse en sentido inverso. Por ejemplo, suponga que estuviera en duda si el factor C es significativo o no. Entonces el experimentador podría preguntar si es razonable esperar que σ pudiera ser tan grande como 10 o más. Si es improbable que σ sea tan grande como 10, entonces puede concluirse que C es significativo.

Se presentan ahora tres ilustrativos ejemplos de diseños factoriales 2^k no replicados.

EJEMPLO 6-3 ··

Transformación de datos en un diseño factorial

Daniel [35b] describe un diseño factorial 2^4 utilizado para estudiar la rapidez de avance de una perforadora como una función de cuatro factores: la carga de la perforadora (A), la rapidez de flujo (B), la velocidad de rotación (C) y el tipo de lodo de perforación usado (D). Los datos del experimento se presentan en la figura 6-17.

En la figura 6-18 se muestra la gráfica de probabilidad normal de las estimaciones de los efectos de este experimento. Con base en esta gráfica, los factores B, C y D, junto con las interacciones BC y BD, requieren interpretación. La figura 6-19 es la gráfica de probabilidad normal de los residuales y la fi-

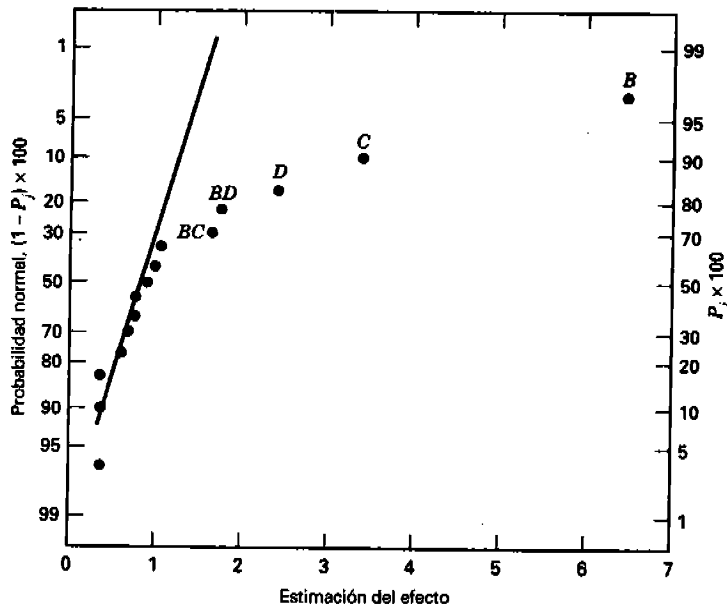


Figura 6-18 Gráfica de probabilidad normal de los efectos del ejemplo 6-3.

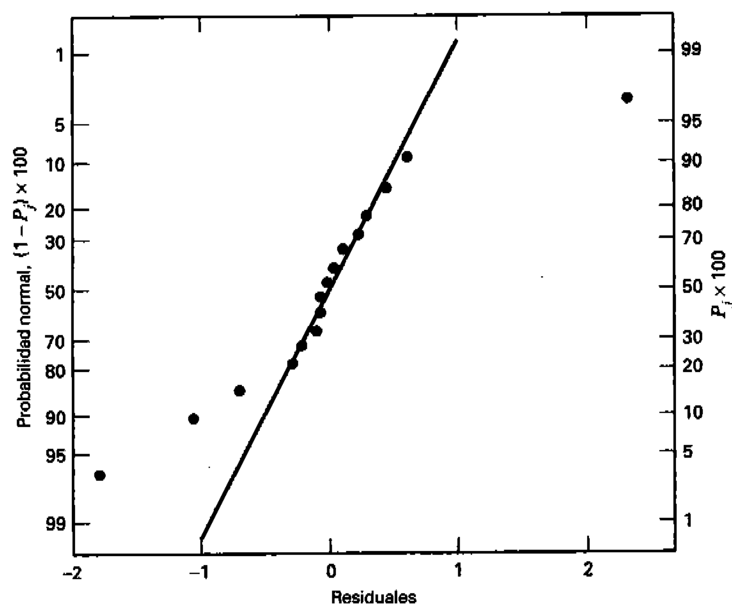


Figura 6-19 Gráfica de probabilidad normal de los residuales del ejemplo 6-3.

gura 6-20 es la gráfica de los residuales contra la velocidad de avance predicha a partir del modelo que contiene los factores identificados. Hay problemas evidentes con la normalidad y la igualdad de la varianza. Con frecuencia se usa una transformación de los datos para abordar estos problemas. Puesto que la variable de respuesta es una razón de cambio, la transformación logarítmica parece un candidato razonable.

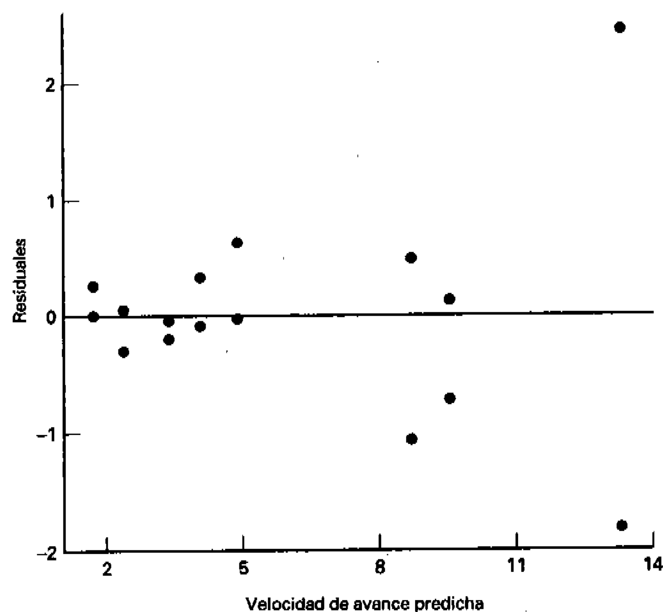


Figura 6-20 Gráfica de los residuales contra la velocidad de avance predicha en el ejemplo 6-3.

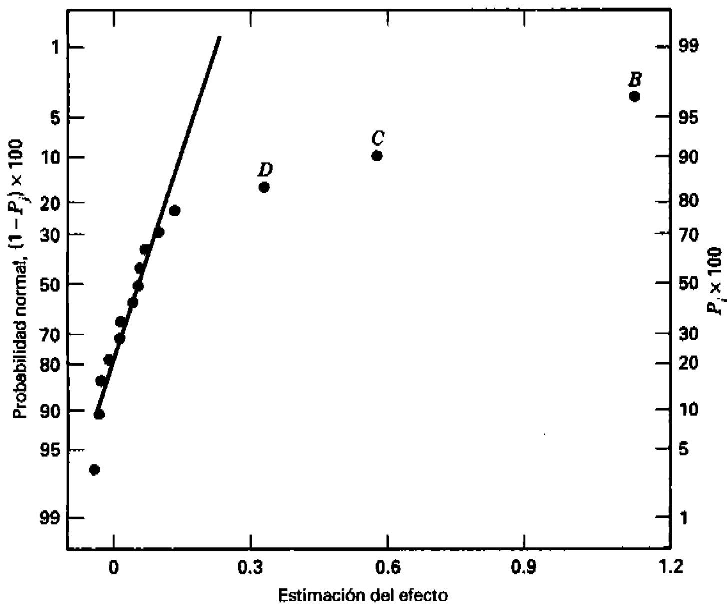


Figura 6-21 Gráfica de probabilidad normal de los efectos del ejemplo 6-3 después de la transformación logarítmica.

En la figura 6-21 se presenta la gráfica de probabilidad normal de las estimaciones de los efectos después de hacer la transformación $y^* = \ln y$. Observe que al parecer ahora es posible una interpretación mucho más simple, ya que sólo los factores B, C y D están activos. Es decir, expresar los datos en la métrica correcta ha simplificado su estructura hasta el punto de que las dos interacciones han dejado de requerirse en el modelo explicatorio.

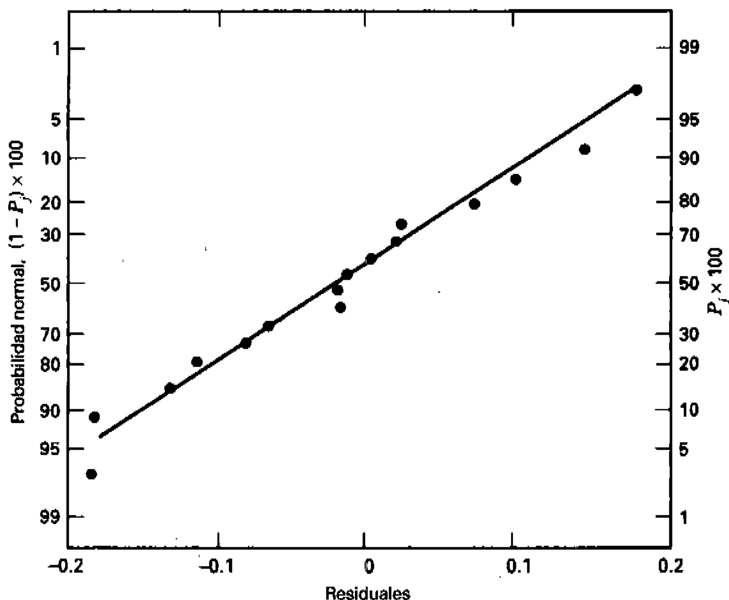


Figura 6-22 Gráfica de probabilidad normal de los residuales del ejemplo 6-3 después de la transformación logarítmica.

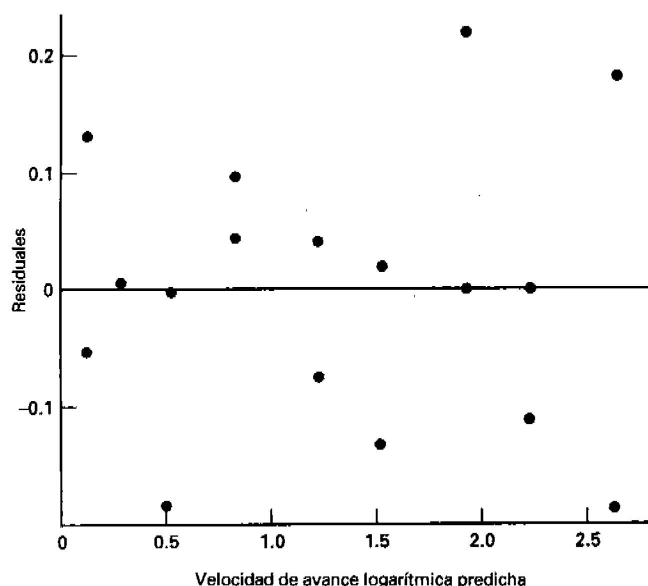


Figura 6-23 Gráfica de los residuales contra la velocidad predicha para el ejemplo 6-3 después de la transformación logarítmica.

En las figuras 6-22 y 6-23 se presentan, respectivamente, una gráfica de probabilidad normal de los residuales y una gráfica de los residuales contra la rapidez de avance predicha para el modelo en la escala logarítmica que contiene a *B*, *C* y *D*. Ahora estas gráficas son satisfactorias. Se concluye que el modelo $y^* = \ln y$ sólo requiere los factores *B*, *C* y *D* para una interpretación adecuada. En la tabla 6-14 se resume el análisis de varianza de este modelo. La suma de cuadrados del modelo es

$$\begin{aligned} SS_{\text{Modelo}} &= SS_B + SS_C + SS_D \\ &= 5.345 + 1.339 + 0.431 \\ &= 7.115 \end{aligned}$$

y $R^2 = SS_{\text{Modelo}}/SS_T = 7.115/7.288 = 0.98$, por lo que el modelo explica cerca de 98% de la variabilidad de la rapidez de avance de la perforadora.

EJEMPLO 6-4

Efectos de localización y dispersión en un diseño factorial no replicado

Se corrió un diseño 2^4 en un proceso de manufactura de paneles laterales y ventanas de un avión comercial. Los paneles se hacen en una prensa, y bajo las condiciones actuales es demasiado elevado el número

Tabla 6-14 Análisis de varianza del ejemplo 6-3 después de la transformación logarítmica

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0	Valor *P*
B (Flujo)	5.345	1	5.345	381.79	<0.0001
C (Velocidad)	1.339	1	1.339	95.64	<0.0001
D (Lodo)	0.431	1	0.431	30.79	<0.0001
Error	0.173	12	0.014		
Total	7.288	15			

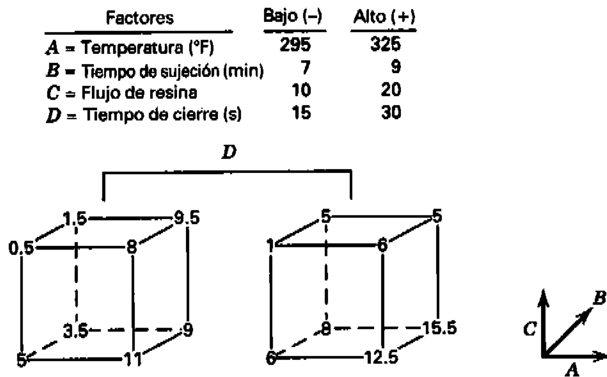


Figura 6-24 Datos del experimento del proceso de los paneles del ejemplo 6-4.

promedio de defectos por panel en una operación de prensado. (El promedio actual del proceso es 5.5 defectos por panel.) Se investigan cuatro factores utilizando una sola réplica de un diseño 2^4, en el que cada réplica corresponde a una sola operación de prensado. Los factores son la temperatura (A), el tiempo de sujeción (B), el flujo de resina (C) y el tiempo de cierre en el prensado (D). En la figura 6-24 se muestran los datos de este experimento.

En la figura 6-25 se muestra la gráfica de probabilidad normal de los efectos de los factores. Es evidente que los dos efectos más grandes son $A = 5.75$ y $C = -4.25$. Ningún efecto de los otros factores parece ser tan grande, y A y C explican cerca de 77% de la variabilidad total, por lo que se concluye que la temperatura (A) baja y el flujo de resina (C) alto reducirían la incidencia de defectos en los paneles.

El análisis residual cuidadoso es un aspecto importante de cualquier experimento. La gráfica de probabilidad normal de los residuales no indicó anomalías, pero cuando el experimentador graficó los resi-

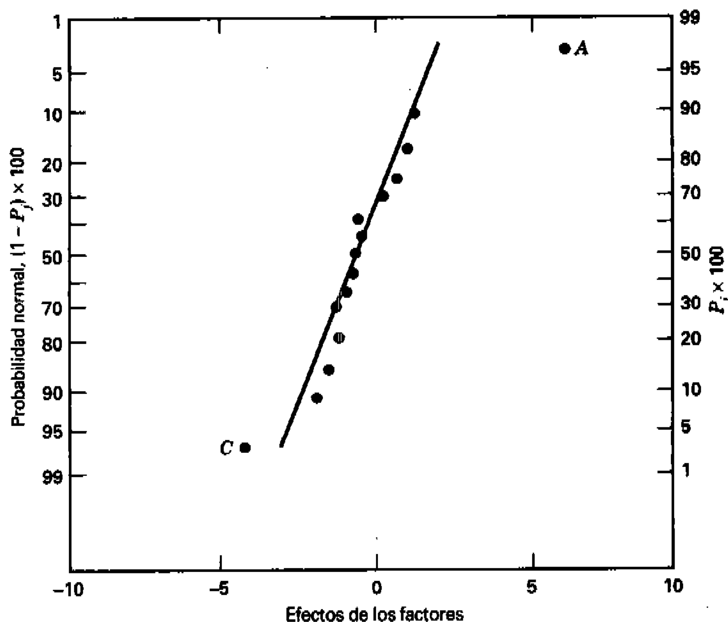


Figura 6-25 Gráfica de probabilidad normal de los efectos de los factores para el experimento del proceso de los paneles del ejemplo 6-4.

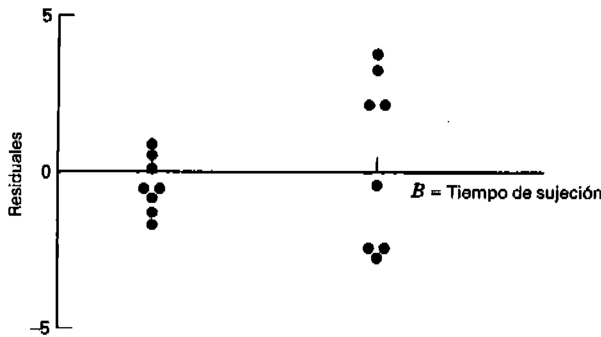


Figura 6-26 Gráfica de los residuales contra el tiempo de sujeción para el ejemplo 6-4.

duales contra cada uno de los factores A a D, la gráfica de los residuales contra B (tiempo de sujeción) presentó el patrón de la figura 6-26. Este factor, que carece de importancia en lo que se refiere al número promedio de defectos por panel, es muy importante en su efecto sobre la variabilidad del proceso, con el tiempo de sujeción bajo dando como resultado una variabilidad menor en el número promedio de defectos por panel en una operación de prensado.

El efecto de dispersión del tiempo de sujeción también es muy evidente en la **gráfica de cubo** de la figura 6-27, donde se grafica el número promedio de defectos por panel y el rango del número de defectos en cada punto del cubo definido por los factores A, B y C. El rango promedio cuando B está en el nivel alto (la cara posterior del cubo de la figura 6-27) es $\overline{R}_{B^+} = 4.75$, y cuando B está en el nivel bajo es $\overline{R}_{B^-} = 1.25$.

Como resultado de este experimento, el ingeniero decidió operar el proceso con la temperatura baja y el flujo de resina alto para reducir el número promedio de defectos, con el tiempo de sujeción bajo para reducir la variabilidad en el número de defectos por panel, y con el tiempo de cierre en el prensado bajo (el cual no tuvo ningún efecto ni sobre la localización ni sobre la dispersión). El nuevo ajuste de las condiciones de operación produjo un nuevo promedio del proceso de menos de un defecto por panel.

Los residuales de un diseño 2^k proporcionan mucha información acerca del problema bajo estudio. Puesto que los residuales pueden considerarse como los valores observados del ruido o error, con frecuencia ofrecen información acerca de la variabilidad del proceso. Puede hacerse el examen sistemático de los residuales de un diseño 2^k no replicado para proporcionar información acerca de la variabilidad del proceso.

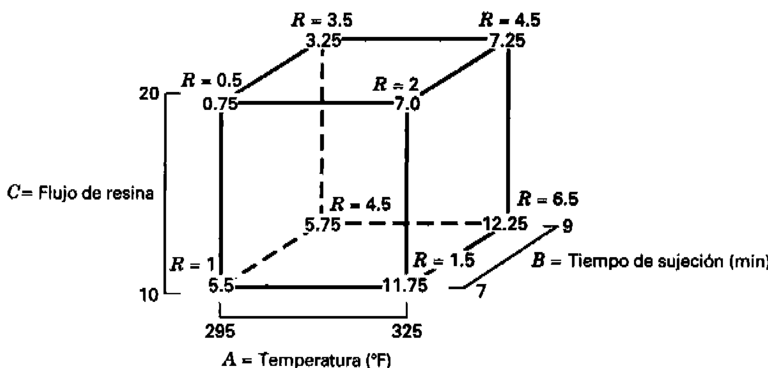


Figura 6-27 Gráfica de cubo de la temperatura, el tiempo de sujeción y el flujo de resina para el ejemplo 6-4.

Tabla 6-15 Cálculo de los efectos de dispersión para el ejemplo 6-4

Corrida	*A*	*B*	*AB*	*C*	*AC*	*BC*	*ABC*	*D*	*AD*	*BD*	*ABD*	*CD*	*ACD*	*BCD*	*ABCD*	Residual
1	−	−	+	−	+	+	−	−	+	+	−	+	−	−	+	−0.94
2	+	−	−	−	−	+	+	−	−	+	+	+	+	−	−	−0.69
3	−	+	−	−	+	−	+	−	+	−	+	+	−	+	−	−2.44
4	+	+	+	−	−	−	−	−	−	−	−	+	+	+	+	−2.69
5	−	−	+	+	−	−	+	−	+	+	−	−	+	+	−	−1.19
6	+	−	−	+	+	−	−	−	−	+	+	−	−	+	+	0.56
7	−	+	−	+	−	+	−	−	+	−	+	−	+	−	+	−0.19
8	+	+	+	+	+	+	+	−	−	−	−	−	−	−	−	2.06
9	−	−	+	−	+	+	−	+	−	−	+	−	+	+	−	0.06
10	+	−	−	−	−	+	+	+	+	−	−	−	−	+	+	0.81
11	−	+	−	−	+	−	+	+	−	+	−	−	+	−	+	2.06
12	+	+	+	−	−	−	−	+	+	+	+	−	−	−	−	3.81
13	−	−	+	+	−	−	+	+	−	−	+	+	−	−	+	−0.69
14	+	−	−	+	+	−	−	+	+	−	−	+	+	−	−	−1.44
15	−	+	−	+	−	+	−	+	−	+	−	+	−	+	−	3.31
16	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	−2.44
$S(i^+)$	2.25	2.72	2.21	1.91	1.81	1.80	1.80	2.24	2.05	2.28	1.97	1.93	1.52	2.09	1.61	
$S(i^-)$	1.85	0.83	1.86	2.20	2.24	2.26	2.24	1.55	1.93	1.61	2.11	1.58	2.16	1.89	2.33	
F_i^*	0.39	2.37	0.34	−0.28	−0.43	−0.46	−0.44	0.74	0.12	0.70	−0.14	0.40	−0.70	0.28	−0.74	

Considere la gráfica de los residuales de la figura 6-26. La desviación estándar de los ocho residuales donde B está en el nivel bajo es $S(B^-) = 0.83$, y la desviación estándar de los ocho residuales donde B está en el nivel alto es $S(B^+) = 2.72$. El estadístico

$$F_B^* = \ln \frac{S^2(B^+)}{S^2(B^-)} \tag{6-24}$$

tiene una distribución aproximadamente normal cuando las dos varianzas $\sigma^2(B^+)$ y $\sigma^2(B^-)$ son iguales. Para ilustrar los cálculos, el valor de F_B^* es

$$\begin{aligned} F_B^* &= \ln \frac{S^2(B^+)}{S^2(B^-)} \\ &= \ln \frac{(2.72)^2}{(0.83)^2} \\ &= 2.37 \end{aligned}$$

En la tabla 6-15 se presenta el conjunto completo de contrastes para el diseño 2^4 junto con los residuales para cada corrida del experimento del proceso de los paneles del ejemplo 6-4. Cada columna de esta tabla contiene el mismo número de signos positivos y negativos, y es posible calcular la desviación estándar de los residuales de cada grupo de signos en cada columna, por ejemplo, $S(i^+)$ y $S(i^-)$, $i = 1, 2, ..., 15$. Entonces

$$F_i^* = \ln \frac{S^2(i^+)}{S^2(i^-)} \qquad i = 1, 2, \ldots, 15 \tag{6-25}$$

es un estadístico que puede usarse para evaluar la magnitud de los **efectos de dispersión** del experimento. Si la varianza de los residuales de las corridas donde el factor i es positivo es igual a la varianza de los residuales de las corridas donde el factor i es negativo, entonces F_i^* tiene una distribución aproximadamente normal. Los valores de F_i^* se presentan al final de cada columna de la tabla 6-15.

La figura 6-28 es la gráfica de probabilidad normal de los efectos de dispersión F_i^*. Evidentemente, B es un factor importante en lo que se refiere a la dispersión del proceso. Para un estudio más amplio de

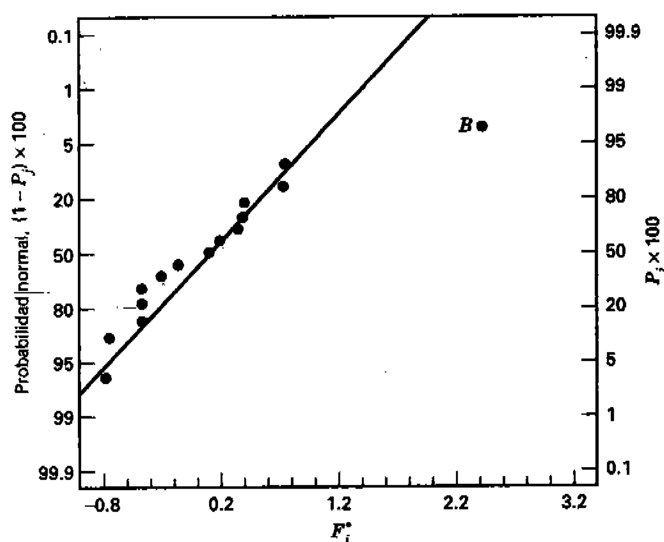


Figura 6-28 Gráfica de probabilidad normal de los efectos de dispersión F_i^* del ejemplo 6-4.

este procedimiento, ver Box y Meyer [19] y Myers y Montgomery [85a]. Asimismo, para que los residuales del modelo ofrezcan la información apropiada acerca de los efectos de dispersión, es necesario especificar correctamente el **modelo de localización**. Referirse al material suplementario del texto de este capítulo para mayores detalles y un ejemplo.

EJEMPLO 6-5 ·······························

Mediciones duplicadas de la respuesta

Un equipo de ingenieros en una fábrica de semiconductores realizaron un diseño factorial 2^4 en un horno de oxidación vertical. Se "apilan" cuatro obleas en el horno, y la variable de respuesta de interés es el espesor del óxido en las obleas. Los cuatro factores del diseño son la temperatura (A), el tiempo (B), la presión (C) y el flujo de gas (D). El experimento se lleva a cabo cargando cuatro obleas en el horno, ajustando las variables del proceso en las condiciones de prueba requeridas por el diseño experimental, procesando las obleas y midiendo después el espesor del óxido en las cuatro obleas. En la tabla 6-16 se presentan el diseño y las mediciones del espesor resultantes. En esta tabla, las cuatro columnas bajo el encabezado "Espesor" contienen las mediciones del espesor del óxido de cada oblea individual, y las dos últimas columnas contienen el promedio muestral y la varianza muestral de las mediciones del espesor en las cuatro obleas de cada corrida.

La manera apropiada de analizar este experimento es considerar las mediciones del espesor de las obleas individuales como **mediciones duplicadas**, y no como réplicas. Si fueran en realidad réplicas, cada oblea se habría procesado individualmente en una sola corrida del horno. Sin embargo, debido a que las cuatro obleas se procesaron en conjunto, recibieron los factores de los tratamientos (es decir, los niveles de las variables del diseño) *simultáneamente*, por lo que hay mucho menos variabilidad en las mediciones del espesor de las obleas individuales que la que se habría observado si cada oblea fuera una réplica. Por lo tanto, el **promedio** de las mediciones del espesor es la variable de respuesta correcta que deberá considerarse inicialmente.

En la tabla 6-17 se muestran las estimaciones de los efectos de este experimento, utilizando el espesor del óxido promedio $\bar{y}$ como la variable de respuesta. Observe que los factores A y B y la interacción AB tienen efectos grandes que explican en conjunto cerca de 90% de la variabilidad del espesor promedio del

Tabla 6-16 El experimento del espesor del óxido

Orden estándar	Orden de la corrida	A	B	C	D	Espesor				$\bar{y}$	s^2
1	10	−1	−1	−1	−1	378	376	379	379	378	2
2	7	1	−1	−1	−1	415	416	416	417	416	0.67
3	3	−1	1	−1	−1	380	379	382	383	381	3.33
4	9	1	1	−1	−1	450	446	449	447	448	3.33
5	6	−1	−1	1	−1	375	371	373	369	372	6.67
6	2	1	−1	1	−1	391	390	388	391	390	2
7	5	−1	1	1	−1	384	385	386	385	385	0.67
8	4	1	1	1	−1	426	433	430	431	430	8.67
9	12	−1	−1	−1	1	381	381	375	383	380	12.00
10	16	1	−1	−1	1	416	420	412	412	415	14.67
11	8	−1	1	−1	1	371	372	371	370	371	0.67
12	1	1	1	−1	1	445	448	443	448	446	6
13	14	−1	−1	1	1	377	377	379	379	378	1.33
14	15	1	−1	1	1	391	391	386	400	392	34
15	11	−1	1	1	1	375	376	376	377	376	0.67
16	13	1	1	1	1	430	430	428	428	429	1.33

Tabla 6-17 Estimaciones de los efectos del ejemplo 6-5, la variable de respuesta es el espesor promedio del óxido

Término del modelo	Estimación del efecto	Suma de cuadrados	Contribución porcentual
A	43.125	7439.06	67.9339
B	18.125	1314.06	12.0001
C	–10.375	430.562	3.93192
D	–1.625	10.5625	0.0964573
AB	16.875	1139.06	10.402
AC	–10.625	451.563	4.12369
AD	1.125	5.0625	0.046231
BC	3.875	60.0625	0.548494
BD	–3.875	60.0625	0.548494
CD	1.125	5.0625	0.046231
ABC	–0.375	0.5625	0.00513678
ABD	2.875	33.0625	0.301929
ACD	–0.125	0.0625	0.000570753
BCD	–0.625	1.5625	0.0142688
ABCD	0.125	0.0625	0.000570753

óxido. La figura 6-29 es una gráfica de probabilidad normal de los efectos. Al examinar esta representación, se concluiría que los factores *A, B* y *C* y las interacciones *AB* y *AC* son importantes. En la tabla 6-18 se muestra el análisis de varianza de este modelo.

El modelo para predecir el espesor promedio del óxido es

$$\hat{y} = 399.19 + 21.56x_1 + 9.06x_2 - 5.19x_3 + 8.44x_1x_2 - 5.31x_1x_3$$

El análisis residual de este modelo es satisfactorio.

Los experimentadores están interesados en obtener un espesor promedio del óxido de 400 Å, y las especificaciones del producto requieren que el espesor sea de entre 390 y 410 Å. En la figura 6-30 se presentan dos gráficas de contorno del espesor promedio, una con el factor *C* (o x_3), la presión, en el nivel bajo (es decir, $x_3 = -1$) y la otra con *C* (o x_3) en el nivel alto (es decir, $x_3 = +1$). Al examinar estas gráficas de contorno, es evidente que hay muchas combinaciones del tiempo y la temperatura (factores *A* y *B*) que producirán resultados aceptables. Sin embargo, si la presión se mantiene constante en el nivel bajo, la "ventana" de operación se corre hacia el extremo izquierdo, o más bajo, del eje del tiempo, indicando que se necesitarán duraciones del ciclo más cortas para conseguir el espesor del óxido deseado.

Es interesante observar los resultados que se hubieran obtenido si las mediciones del espesor del óxido de las obleas se hubieran considerado **incorrectamente** como réplicas. En la tabla 6-19 se presenta el análisis de varianza del modelo completo basado en tratar el experimento como un diseño factorial 2^4 con réplicas. Observe que hay muchos factores significativos en este análisis, lo cual sugiere un modelo mucho más complejo del que se encontró cuando se utilizó el espesor promedio del óxido como la respuesta. La razón de esto es que la estimación de la varianza del error de la tabla 6-19 es muy pequeña ($\hat{\sigma}^2 = 6.12$). El cuadrado medio de los residuales de la tabla 6-19 refleja la variabilidad entre las obleas *dentro* de una corrida y la variabilidad *entre* las corridas. La estimación del error que se obtiene en la tabla 6-18 es mucho más grande, $\hat{\sigma}^2 = 17.61$, y es principalmente una medida de la variabilidad entre las corridas. Ésta es la mejor estimación del error que deberá usarse para juzgar la significación de las variables del proceso que se modifican de una corrida a otra.

Una pregunta lógica que podría plantearse es: ¿qué daño causa identificar demasiados factores como importantes?, como ciertamente sería el caso en el análisis incorrecto de la tabla 6-19. La respuesta es que

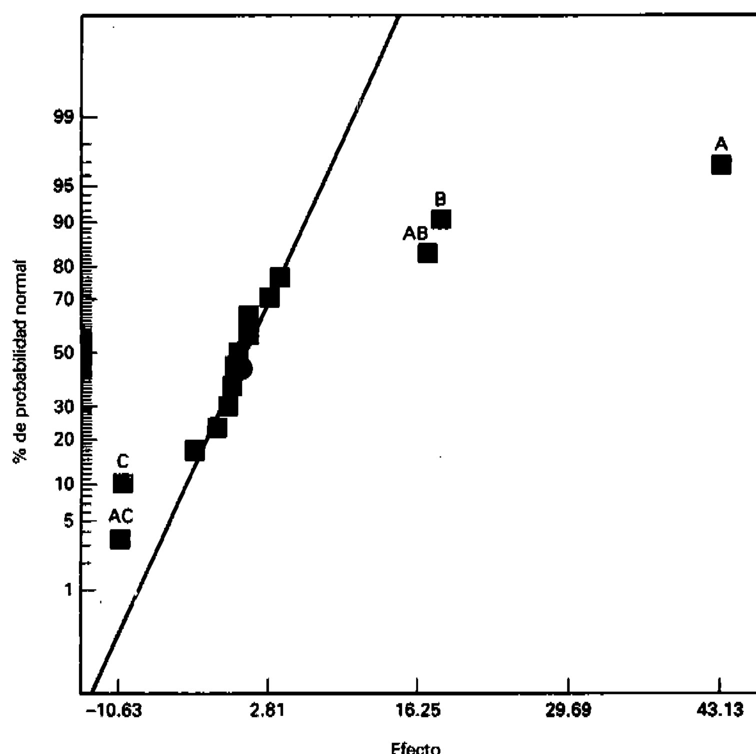


Figura 6-29 Gráfica de probabilidad normal de los efectos para la respuesta del espesor promedio del óxido, ejemplo 6-5.

Tabla 6-18 Análisis de varianza (de *Design-Expert*) para la respuesta espesor promedio del óxido, ejemplo 6-5

Source	Sum of Squares	DF	Mean Square	F Value	Prob > F
Model	10774.31	5	2154.86	122.35	<0.000
A	*7439.06*	*1*	*7439.06*	*422.37*	*<0.000*
B	*1314.06*	*1*	*1314.06*	*74.61*	*<0.000*
C	*430.56*	*1*	*430.56*	*24.45*	*0.0006*
AB	*1139.06*	*1*	*1139.06*	*64.67*	*<0.000*
AC	*451.56*	*1*	*451.56*	*25.64*	*0.0005*
Residual	176.12	10	17.61		
Cor Total	10950.44	15			

Std. Dev.	4.20	*R*-Squared	0.9839
Mean	399.19	Adj. *R*-Squared	0.9759
C.V.	1.05	Pred. *R*-Squared	0.9588
PRESS	450.88	Adeq. Precision	27.967

Factor	Coefficient Estimate	DF	Standard Error	95% CI Low	95% CI High
Intercept	399.19	1	1.05	396.85	401.53
A-Time	21.56	1	1.05	19.22	23.90
B-Temp	9.06	1	1.05	6.72	11.40
C-Pressure	−5.19	1	1.05	−7.53	−2.85
AB	8.44	1	1.05	6.10	10.78
AC	−5.31	1	1,05	−7.65	−2.97

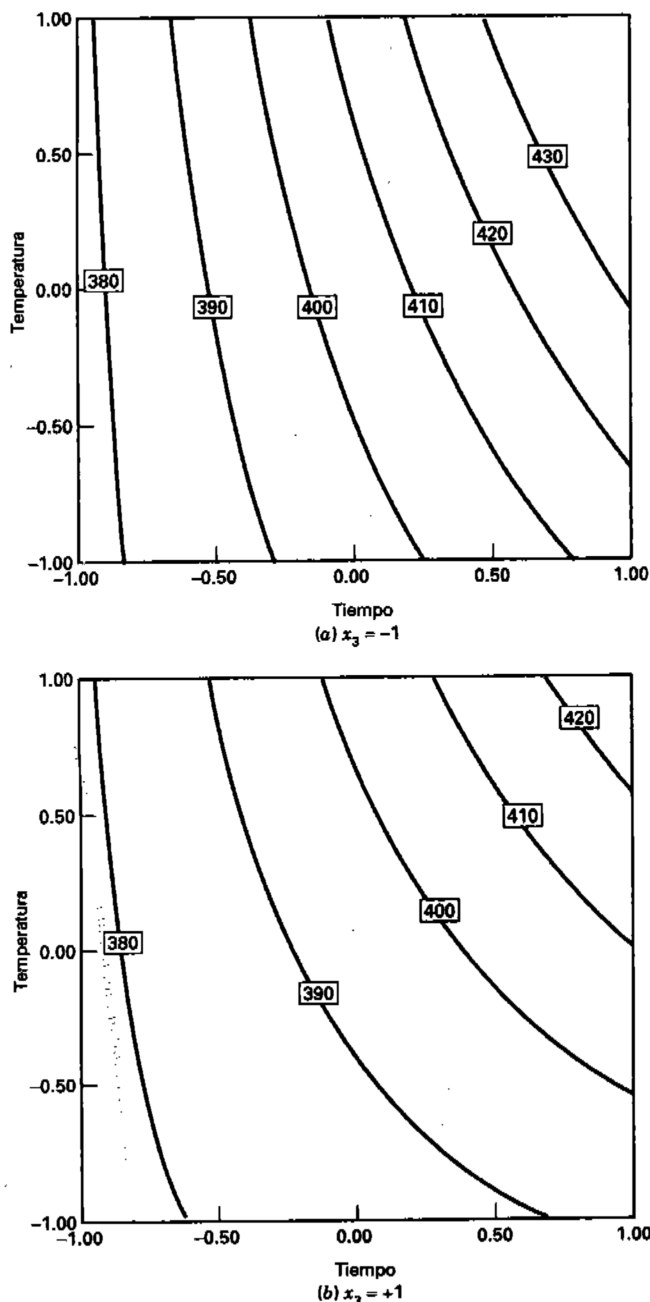


Figura 6-30 Gráficas de contorno del espesor promedio del óxido con la presión (x_3) mantenida constante.

intentar manipular u optimizar los factores que no son importantes sería un desperdicio de recursos, y podría resultar en agregar variabilidad innecesaria a *otras* respuestas de interés.

Cuando se hacen mediciones duplicadas de la respuesta, casi siempre hay información útil acerca de algún aspecto de la variabilidad del proceso contenida en estas observaciones. Por ejemplo, si las mediciones duplicadas son pruebas múltiples hechas con un instrumento de medición en la misma unidad ex-

Tabla 6-19 Análisis de varianza (de *Design-Expert*) de la respuesta individual del espesor del óxido de las obleas

Source	Sum of Squares	DF	Mean Square	F Value	Prob > F
Model	43801.75	15	2920.12	476.75	<0.0001
A	*29756.25*	*1*	*29756.25*	*4858.16*	*<0.0001*
B	*5256.25*	*1*	*5256.25*	*858.16*	*<0.0001*
C	*1722.25*	*1*	*1722.25*	*281.18*	*<0.0001*
D	*42.25*	*1*	*42.25*	*6.90*	*0.0115*
AB	*4556.25*	*1*	*4556.25*	*743.88*	*<0.0001*
AC	*1806.25*	*1*	*1806.25*	*294.90*	*<0.0001*
AD	*20.25*	*1*	*20.25*	*3.31*	*0.0753*
BC	*240.25*	*1*	*240.25*	*39.22*	*<0.0001*
BD	*240.25*	*1*	*240.25*	*39.22*	*<0.0001*
CD	*20.25*	*1*	*20.25*	*3.31*	*0.0753*
ABD	*132.25*	*1*	*132.25*	*21.59*	*<0.0001*
ABC	*2.25*	*1*	*2.25*	*0.37*	*0.5473*
ACD	*0.25*	*1*	*0.25*	*0.041*	*0.8407*
BCD	*6.25*	*1*	*6.25*	*1.02*	*0.3175*
ABCD	*0.25*	*1*	*0.25*	*0.041*	*0.8407*
Residual	294.00	48	6.12		
Lack of Fit	*0.000*	*0*			
Pure Error	*294.00*	*48*	*6.13*		
Cor. Total	44095.75	63			

perimental, entonces las mediciones duplicadas proporcionan cierta información acerca de la eficiencia del instrumento de medición. Si las mediciones duplicadas se hacen en diferentes lugares dentro de una unidad experimental, pueden brindar cierta información acerca de la *uniformidad* de la variable de respuesta en esa unidad. En el ejemplo tratado aquí, ya que se tiene una observación en cada una de cuatro unidades experimentales que se han sometido a un procesamiento conjunto, se tiene cierta información acerca de la variabilidad *dentro de las corridas* del proceso. Esta información se encuentra contenida en la varianza de las mediciones del espesor del óxido de las cuatro obleas de cada corrida. Sería de interés determinar si alguna de las variables del proceso influye en la variabilidad al interior de las corridas.

La figura 6-31 es una gráfica de probabilidad normal de las estimaciones de los efectos obtenidas utilizando $\ln(s^2)$ como la respuesta. Recuerde que en el capítulo 3 se indicó que la transformación logarítmica es por lo general apropiada para modelar la variabilidad. No hay ningún efecto individual fuerte, pero el factor A y la interacción BD son los más grandes. Si se incluyen también los efectos principales de B y D para obtener un modelo jerárquico, entonces el modelo de $\ln(s^2)$ es

$$\widehat{\ln(s^2)} = 1.08 + 0.41x_1 - 0.40x_2 + 0.20x_4 - 0.56x_2x_4$$

El modelo explica apenas poco menos de la mitad de la variabilidad en la respuesta $\ln(s^2)$, lo cual desde luego no es nada espectacular para un modelo empírico, pero con frecuencia es difícil obtener modelos excepcionalmente buenos de las varianzas.

La figura 6-32 es una gráfica de contorno de la varianza predicha (no del logaritmo de la varianza predicha) con la presión x_3 en el nivel bajo (recuerde que con esto se minimiza la duración del ciclo) y el flujo de gas x_4 en el nivel alto. Esta elección del flujo de gas produce los valores mínimos de la varianza predicha en la región de la gráfica de contorno.

En este caso, los experimentadores se enfocaron en seleccionar valores de las variables de diseño que dieran un espesor medio del óxido dentro de las especificaciones del proceso y tan cerca de 400 Å como fuera posible, haciendo al mismo tiempo que la variabilidad dentro de las corridas sea pequeña, por ejemplo $s^2 \leq 2$.

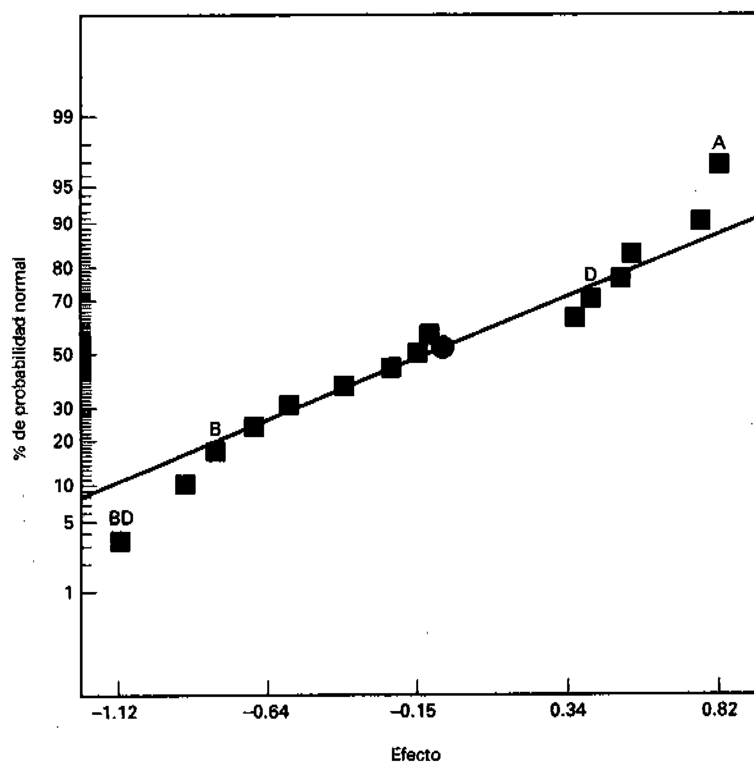


Figura 6-31 Gráfica de probabilidad normal de los efectos utilizando ln (s^2) como la respuesta, ejemplo 6-5.

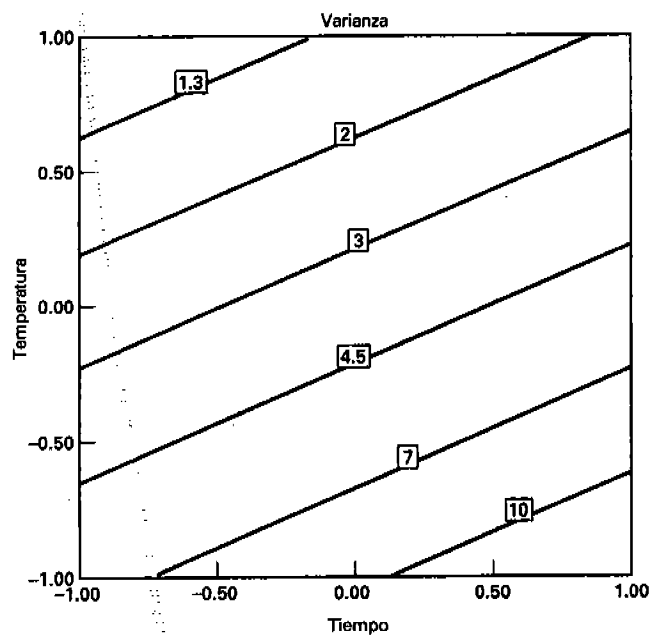


Figura 6-32 Gráfica de contorno de s^2 (variabilidad dentro de las corridas) con la presión en el nivel bajo y el flujo de gas en el nivel alto.

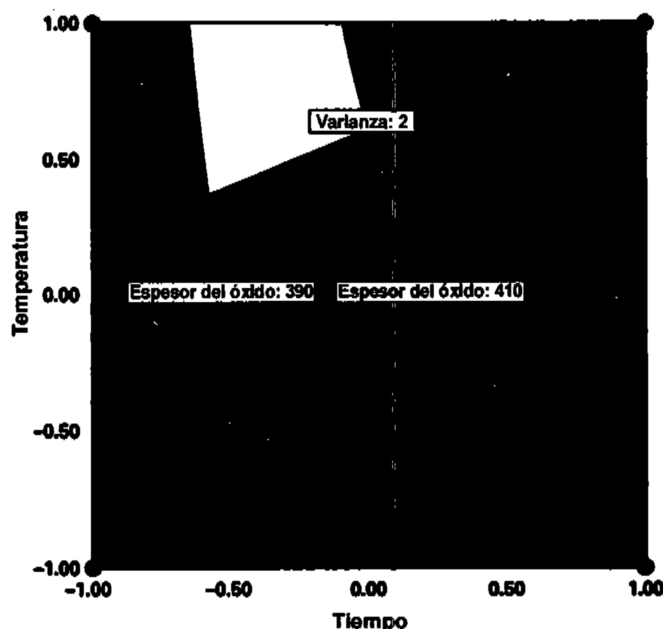


Figura 6-33 Superposición del espesor promedio del óxido y las respuestas s^2 con la presión en el nivel bajo y el flujo de gas en el nivel alto.

Una manera posible de encontrar un conjunto de condiciones adecuado es superponiendo las gráficas de contorno de las figuras 6-30 y 6-32. La gráfica de la superposición se muestra en la figura 6-33, con las especificaciones del espesor medio del óxido y la restricción $s^2 \leq 2$ indicadas como contornos. En esta gráfica, la presión se mantiene constante en el nivel bajo y el flujo de gas se mantiene constante en el nivel alto. La región no sombreada cerca de la parte central izquierda de la gráfica identifica una región factible para las variables tiempo y temperatura.

Éste es un ejemplo simple del uso de las gráficas de contorno para estudiar dos respuestas simultáneamente. Este problema se analizará con mayor detalle en el capítulo 11.

6-6 ADICIÓN DE PUNTOS CENTRALES EN EL DISEÑO 2^k

Una preocupación potencial en el uso de diseños factoriales de dos niveles es el supuesto de la **linealidad** de los efectos de los factores. Desde luego, no es necesaria la linealidad perfecta, y el sistema 2^k funcionará bastante bien incluso cuando el supuesto de linealidad sea válido sólo de manera muy aproximada. De hecho, se ha señalado ya que si se agregan los **términos de interacción** a un modelo de los efectos principales o de primer orden, de donde se obtiene

$$y=\beta_0+\sum_{j=1}^{k}\beta_j x_j+\sum_{i<j}\sum\beta_{ij}x_i x_j+\varepsilon \tag{6-26}$$

entonces se tiene un modelo con la capacidad de representar cierta curvatura en la función de respuesta. Esta curvatura, desde luego, es resultado del torcimiento del plano inducido por los términos de interacción $\beta_{ij}x_ix_j$.

Habrá situaciones en que la curvatura de la función de respuesta no esté modelada adecuadamente por la ecuación 6-26. En tales casos, un modelo lógico por considerar es

$$y = \beta_0 + \sum_{j=1}^{k} \beta_j x_j + \sum_{i<j}\sum \beta_{ij} x_i x_j + \sum_{j=1}^{k} \beta_{jj} x_j^2 + \varepsilon \qquad (6\text{-}27)$$

donde las β_{jj} representan **efectos cuadráticos** o de segundo orden puros. A la ecuación 6-27 se le llama **modelo de superficie de respuesta de segundo orden.**

Cuando se realiza un experimento factorial de dos niveles, por lo general se anticipa el ajuste del modelo de primer orden de la ecuación 6-26, pero deberá estarse alerta ante la posibilidad de que el modelo de segundo orden de la ecuación 6-27 sea en realidad más apropiado. Existe un método para hacer una réplica de ciertos puntos de un diseño factorial 2^k que ofrecerá protección contra la curvatura de los efectos de segundo orden a la vez que permitirá una estimación independiente del error que va a obtenerse. El método consiste en agregar **puntos centrales** en el diseño 2^k. Éstos consisten en n réplicas que se corren en los puntos $x_i = 0$ $(i = 1, 2, ..., k)$. Una razón importante para agregar réplicas de las corridas en el centro del diseño es que los puntos centrales no afectan las estimaciones usuales de los efectos en un diseño 2^k. Cuando se agregan puntos centrales, se supone que los k factores son **cuantitativos.**

Para ilustrar este enfoque, considere un diseño 2^2 con una observación en cada uno de los puntos factoriales (–, –), (+, –), (–, +) y (+, +), y n_C observaciones en el punto central (0, 0). En la figura 6-34 se ilustra la situación. Sea $\bar{y}_F$ el promedio de las cuatro corridas en los cuatro puntos factoriales y sea $\bar{y}_C$ el promedio de las n_C corridas en el punto central. Si la diferencia $\bar{y}_F - \bar{y}_C$ es pequeña, entonces los puntos centrales caen en el plano (o cerca de él) que pasa por los puntos factoriales, y no hay curvatura cuadrática. Por otra parte, si $\bar{y}_F - \bar{y}_C$ es grande, entonces está presente una curvatura cuadrática. La **suma de cuadrados de la curvatura cuadrática pura** con un solo grado de libertad está dada por

$$SS_{\text{Cuadrática pura}} = \frac{n_F n_C (\bar{y}_F - \bar{y}_C)^2}{n_F + n_C} \qquad (6\text{-}28)$$

donde, en general, n_F es el número de puntos del diseño factorial. Esta cantidad puede compararse con el cuadrado medio del error para probar la curvatura cuadrática pura. Más específicamente, cuando se

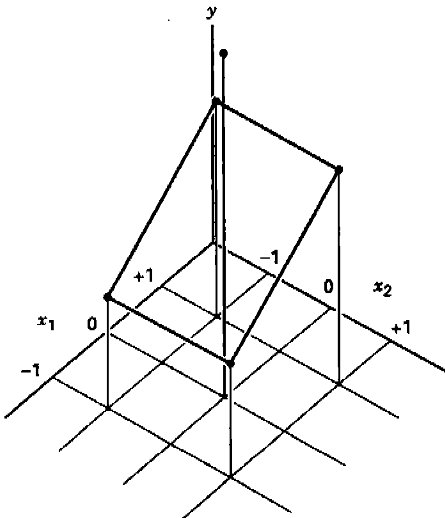


Figura 6-34 Diseño 2^2 con puntos centrales.

agregan puntos en el centro del diseño 2^k, con la prueba de la curvatura (utilizando la ecuación 6-28) en realidad se prueban las hipótesis

$$H_0: \sum_{j=1}^{k} \beta_{jj} = 0$$

$$H_1: \sum_{j=1}^{k} \beta_{jj} \neq 0$$

Además, si los puntos factoriales del diseño no tienen réplicas, pueden usarse los n_C puntos centrales para construir una estimación del error con $n_C - 1$ grados de libertad.

EJEMPLO 6-6 ··

Un ingeniero químico estudia el rendimiento de un proceso. Hay dos variables de interés, el tiempo de reacción y la temperatura de reacción. Debido a que no se tiene la seguridad sobre el supuesto de linealidad en la región de exploración, el ingeniero decide realizar un diseño factorial 2^2 (con una sola réplica de cada corrida factorial) aumentando con cinco puntos centrales. El diseño y los datos del rendimiento se muestran en la figura 6-35.

En la tabla 6-20 se resume el análisis de varianza de este experimento. El cuadrado medio del error se calcula a partir de los puntos centrales de la siguiente manera:

$$MS_E = \frac{SS_E}{n_C - 1} = \frac{\sum_{\text{Puntos centrales}} (y_i - \bar{y})^2}{n_C - 1} \tag{6-29}$$

Por lo tanto, por la tabla 6-20,

$$MS_E = \frac{\sum_{i=1}^{5} (y_i - 40.46)^2}{4} = \frac{0.1720}{4} = 0.0430$$

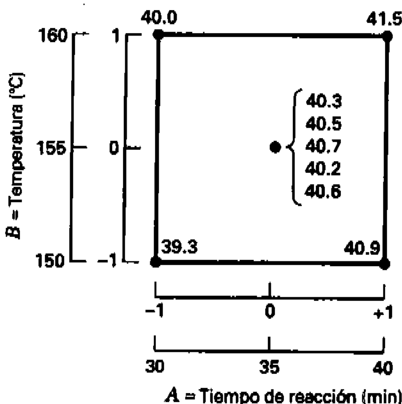


Figura 6-35 El diseño 2^2 con cinco puntos centrales para el ejemplo 6-6.

Tabla 6-20 Análisis de varianza del ejemplo 6-6

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0	Valor P
A (Tiempo)	2.4025	1	2.4025	55.87	0.0017
B (Temperatura)	0.4225	1	0.4225	9.83	0.0350
AB	0.0025	1	0.0025	0.06	0.8185
Cuadrática pura	0.0027	1	0.0027	0.06	0.8185
Error	0.1720	4	0.0430		
Total	3.0022	8			

El promedio de los puntos de la parte factorial del diseño es $\bar{y}_F = 40.425$, y el promedio de los puntos situados en el centro es $\bar{y}_C = 40.46$. La diferencia $\bar{y}_F - \bar{y}_C = 40.425 - 40.46 = -0.035$ parece ser pequeña. La suma de cuadrados de la curvatura cuadrática pura de la tabla del análisis de varianza se calcula con la ecuación 6-28 de la siguiente manera:

$$
\begin{aligned}
SS_{\text{Cuadrática pura}} &= \frac{n_F n_C (\bar{y}_F - \bar{y}_C)^2}{n_F + n_C} \\
&= \frac{(4)(5)(-0.035)^2}{4+5} \\
&= 0.0027
\end{aligned}
$$

El análisis de varianza indica que ambos factores tienen efectos principales significativos, que no existe interacción, y que no hay evidencia de curvatura de segundo orden en la respuesta en la región de exploración. Es decir, la hipótesis nula H_0: $\beta_{11} + \beta_{22} = 0$ no puede rechazarse.

En el ejemplo 6-6 se llegó a la conclusión de que no había indicios de efectos cuadráticos; es decir, un modelo de primer orden

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \varepsilon$$

es apropiado (aun cuando probablemente no se necesite el término de la interacción). Habrá situaciones en las que se necesitarán los términos cuadráticos. Es decir, se tendrá que suponer entonces un modelo de segundo orden tal como

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \varepsilon$$

Desafortunadamente, los parámetros desconocidos (las β) de este modelo no pueden estimarse, ya que hay seis parámetros por estimar y el diseño 2^2 más los puntos centrales de la figura 6-35 sólo tienen cinco corridas independientes.

Una solución simple y de gran efectividad de este problema es aumentar el diseño 2^2 con cuatro **corridas axiales,** como se ilustra en la figura 6-36*a*. El diseño resultante, llamado **diseño central compuesto,** puede usarse entonces para ajustar el modelo de segundo orden. En la figura 6-36*b* se muestra un diseño central compuesto para $k = 3$ factores. Este diseño tiene $14 + n_C$ corridas (generalmente $3 \leq n_C \leq 5$), y es un diseño muy eficiente para ajustar el modelo de segundo orden con 10 parámetros en $k = 3$ factores.

Los diseños compuestos centrales se usan ampliamente para construir modelos de superficie de respuesta de segundo orden. Estos diseños se estudiarán con mayor detalle en el capítulo 11.

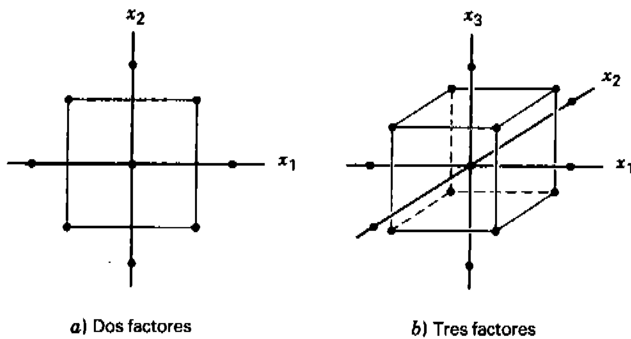


Figura 6-36 Diseños centrales compuestos.

Se concluye esta sección con algunas sugerencias y observaciones adicionales útiles referentes al uso de puntos centrales.

1. Cuando un experimento factorial se lleva a cabo en un proceso en marcha, considere utilizar las condiciones de operación actuales (o de receta) como el punto central del diseño. Esto con frecuencia le asegura al personal de operación que al menos una parte de las corridas del experimento van a realizarse bajo condiciones familiares, y por lo tanto es improbable que los resultados obtenidos (por lo menos para estas corridas) sean peores que los que se obtienen típicamente.
2. Cuando el punto central de un experimento factorial corresponde con las condiciones de operación actuales, el experimentador puede usar las respuestas observadas en el punto central para proporcionar una verificación aproximada de si algo "inusual" ocurrió durante el experimento. Es decir, las respuestas del punto central deberán ser muy similares a las respuestas observadas históricamente en la operación rutinaria del proceso. Con frecuencia el personal de operación llevará una carta de control para monitorear el desempeño del proceso. En ocasiones las respuestas de los puntos centrales pueden graficarse directamente en la carta de control como una verificación de la forma en que estuvo operando el proceso durante el experimento.
3. Considere correr las réplicas del punto central en orden no aleatorio. Específicamente, deberán correrse uno o dos puntos centrales en o cerca del principio del experimento, uno o dos cerca de la parte media, y uno o dos cerca del final. Al separar los puntos centrales en el tiempo, el experimentador tiene una verificación aproximada de la estabilidad del proceso durante el experimento. Por ejemplo, si ha ocurrido una tendencia en la respuesta mientras se realizaba el experimento, graficar las respuestas de los puntos centrales contra el tiempo puede poner de manifiesto esta situación.
4. En ocasiones los experimentos tienen que realizarse en situaciones en las que la información previa acerca de la variabilidad del proceso es escasa o nula. En estos casos, correr dos o tres puntos centrales como las primeras corridas en el experimento puede ser de suma utilidad. Estas corridas pueden proporcionar una estimación preliminar de la variabilidad. Si la magnitud de la variabilidad parece razonable, se continúa; por otra parte, si la variabilidad observada es mayor que la anticipada (¡o que la razonable!), habrá que detenerse. Con frecuencia es muy provechoso estudiar la cuestión de por qué es tan grande la variabilidad antes de proceder con el resto del experimento.
5. Generalmente, se utilizan puntos centrales cuando todos los factores del diseño son **cuantitativos.** Sin embargo, en ocasiones habrá una o más variables cualitativas o categóricas y varias cuan-

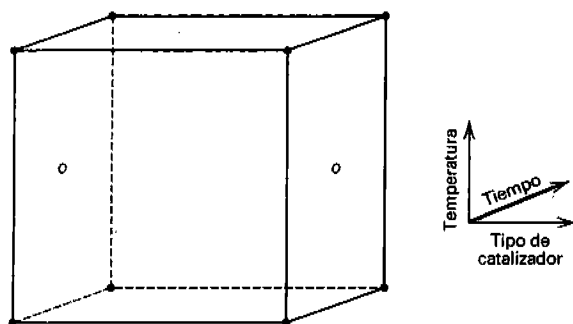


Figura 6-37 Un diseño 2^3 con un factor cualitativo y puntos centrales.

titativas. Sigue siendo posible emplear los puntos centrales en estos casos. Para ilustrar este punto, considere un experimento con dos factores cuantitativos, el tiempo y la temperatura, cada uno con dos niveles, y un solo factor cualitativo, el tipo de catalizador, también con dos niveles (orgánico e inorgánico). En la figura 6-37 se muestra el diseño 2^3 para estos factores. Observe que los puntos centrales se colocan en las caras opuestas del cubo que incluyen los factores cuantitativos. En otras palabras, los puntos centrales pueden correrse con las combinaciones de los tratamientos en los niveles alto y bajo de los factores cualitativos, siempre y cuando esos subespacios incluyan únicamente factores cuantitativos.

6-7 PROBLEMAS

6-1. Un ingeniero está interesado en los efectos de la velocidad de corte (A), la geometría de la herramienta (B) y el ángulo de corte (C) sobre la vida (en horas) de una máquina herramienta. Se eligen dos niveles de cada factor y se corren tres réplicas de un diseño factorial 2^3. Los resultados fueron los siguientes:

			Combinación de	Réplica		
A	*B*	*C*	tratamientos	I	II	III
−	−	−	(1)	22	31	25
+	−	−	*a*	32	43	29
−	+	−	*b*	35	34	50
+	+	−	*ab*	55	47	46
−	−	+	*c*	44	45	38
+	−	+	*ac*	40	37	36
−	+	+	*bc*	60	50	54
+	+	+	*abc*	39	41	47

a) Estimar los efectos de los factores. ¿Qué efectos parecen ser grandes?
b) Usar el análisis de varianza para confirmar las conclusiones del inciso *a*.
c) Escribir un modelo de regresión para predecir la vida de la herramienta (en horas) con base en los resultados de este experimento.
d) Analizar los residuales. ¿Hay algún problema evidente?
e) Con base en el análisis de las gráficas de los efectos principales y las interacciones, ¿cuáles serían los niveles de *A*, *B* y *C* que se recomendaría utilizar?

6-2. Considere nuevamente el inciso *c* del problema 6-1. Utilizar el modelo de regresión para generar las gráficas de la superficie de respuesta y de contorno de la respuesta, la vida de la herramienta. Interpretar estas gráficas. ¿Ofrecen alguna idea respecto de las condiciones de operación deseables para este proceso?

6-3. Encontrar el error estándar de los efectos de los factores y aproximar los límites de confianza de 95% para los efectos de los factores en el problema 6-1. ¿Los resultados de este análisis concuerdan con las conclusiones del análisis de varianza?

6-4. Representar los efectos de los factores del problema 6-1 en una gráfica relativa a una distribución t escalada apropiadamente. ¿En esta representación gráfica se identifican de manera adecuada los factores importantes? Comparar las conclusiones de esta gráfica con los resultados del análisis de varianza.

6-5. Se usa una máquina para hacer ranuras de localización en una tarjeta de circuitos impresos. El nivel de vibración en la superficie de la tarjeta cuando se hacen las ranuras se considera una fuente principal de variación dimensional de las ranuras. Se piensa que dos factores influyen en la vibración: el tamaño de las ranuras (A) y la velocidad de corte (B). Se seleccionan dos tamaños de las ranuras ($\frac{1}{16}$ y $\frac{1}{8}$ de pulgada) y dos velocidades (40 y 90 rpm), y se hacen ranuras en cuatro tarjetas con cada conjunto de condiciones que se muestran abajo. La variable de respuesta es la vibración medida como el vector resultante de tres acelerómetros (x, y y z) en cada tarjeta de prueba.

A	B	Combinación de tratamientos	Réplica I	II	III	IV
−	−	(1)	18.2	18.9	12.9	14.4
+	−	*a*	27.2	24.0	22.4	22.5
−	+	*b*	15.9	14.5	15.1	14.2
+	+	*ab*	41.0	43.9	36.3	39.9

a) Analizar los datos de este experimento.

b) Construir una gráfica de probabilidad normal de los residuales, y graficar los residuales contra el nivel de vibración predicho. Interpretar estas gráficas.

c) Hacer la gráfica de la interacción AB. Interpretar esta gráfica. ¿Qué niveles del tamaño de las ranuras y la velocidad se recomendarían para la operación rutinaria?

6-6. Considere nuevamente el experimento descrito en el problema 6-1. Suponga que el experimentador efectuó únicamente ocho ensayos de la réplica 1. Además, corrió cuatro puntos centrales y obtuvo los siguientes valores de la respuesta: 36, 40, 43, 45.

a) Estimar los efectos de los factores. ¿Qué efectos son grandes?

b) Efectuar un análisis de varianza, incluyendo una verificación de la curvatura cuadrática pura. ¿A qué conclusiones se llega?

c) Escribir un modelo apropiado para predecir la vida de la herramienta, con base en los resultados de este experimento. ¿Este modelo difiere en alguna forma sustancial del modelo del problema 6-1, inciso *c*?

d) Analizar los residuales.

e) ¿A qué conclusiones se llegaría acerca de las condiciones de operación apropiadas para este proceso?

6-7. Se llevó a cabo un experimento para mejorar el rendimiento de un proceso químico. Se seleccionaron cuatro factores y se corrieron dos réplicas de un experimento completamente aleatorizado. Los resultados se presentan en la tabla siguiente:

Combinación de tratamientos	Réplica I	II	Combinación de tratamientos	Réplica I	II
(1)	90	93	*d*	98	95
a	74	78	*ad*	72	76
b	81	85	*bd*	87	83
ab	83	80	*abd*	85	86
c	77	78	*cd*	99	90
ac	81	80	*acd*	79	75
bc	88	82	*bcd*	87	84
abc	73	70	*abcd*	80	80

a) Estimar los efectos de los factores.

b) Construir la tabla del análisis de varianza y determinar cuáles factores son importantes para explicar el rendimiento.

c) Escribir un modelo de regresión para predecir el rendimiento, suponiendo que los cuatro factores se hicieron variar en el rango de −1 a +1 (en unidades codificadas).

d) Graficar los residuales contra el rendimiento predicho y en una escala de probabilidad normal. ¿El análisis residual parece ser satisfactorio?

e) Dos interacciones de tres factores, *ABC* y *ABD*, aparentemente tienen efectos grandes. Trazar una gráfica de cubo en los factores *A*, *B* y *C* con los rendimientos promedio indicados en cada vértice. Repetir lo anterior utilizando los factores *A*, *B* y *D*. ¿Estas dos gráficas ayudan en la interpretación de los datos? ¿Dónde se recomendaría que se corriera el proceso con respecto a las cuatro variables?

6-8. Un bacteriólogo está interesado en los efectos de dos medios de cultivo diferentes y dos tiempos diferentes sobre el crecimiento de un virus particular. Realiza seis réplicas de un diseño 2^2, haciendo las corridas de manera aleatoria. Analizar los datos del crecimiento viral que se presentan enseguida y sacar las conclusiones apropiadas. Analizar los residuales y comentar la adecuación del modelo.

	Medio de cultivo			
Tiempo, h	1		2	
	21	22	25	26
12	23	28	24	25
	20	26	29	27
	37	39	31	34
18	38	38	29	33
	35	36	30	35

6-9. Un ingeniero industrial empleado por una compañía refresquera está interesado en los efectos de dos diferentes tipos de botellas de 32 onzas sobre el tiempo de entrega de cajas de 12 botellas del producto. Los dos tipos de botellas son de vidrio y de plástico. Se usan dos empleados para realizar una tarea que consiste en mover 40 cajas del producto 50 pies en una plataforma de carga estándar y acomodarlas en un estante de venta. Se hacen cuatro réplicas de un diseño factorial 2^2 y los tiempos observados se enlistan en la siguiente tabla. Analizar los datos y sacar las conclusiones apropiadas. Analizar los residuales y comentar la adecuación del modelo.

	Empleado			
Tipo de botella	1		2	
Vidrio	5.12	4.89	6.65	6.24
	4.98	5.00	5.49	5.55
Plástico	4.95	4.95	5.28	4.91
	4.27	4.25	4.75	4.71

6-10. En el problema 6-9, el ingeniero también estuvo interesado en las diferencias en la fatiga potencial que resulta de los tipos de botellas. Como una medida de la cantidad de esfuerzo requerido, midió el aumento del ritmo cardiaco (pulso) inducido por la tarea. Los resultados se presentan a continuación. Analizar los datos y sacar conclusiones. Analizar los residuales y comentar la adecuación del modelo.

	Empleado			
Tipo de botella	1		2	
Vidrio	39	45	20	13
	58	35	16	11
Plástico	44	35	13	10
	42	21	16	15

6-11. Calcular los límites de confianza aproximados para los efectos de los factores del problema 6-10. ¿Los resultados de este análisis concuerdan con el análisis de varianza realizado en el problema 6-10?

6-12. En un artículo de *AT&T Technical Journal* (vol. 65, pp. 39-50) se describe la aplicación de diseños factoriales de dos niveles en la fabricación de circuitos integrados. Un paso básico del procesamiento es hacer crecer una capa epitaxial sobre obleas de silicio pulidas. Las obleas se montan en un susceptor, se colocan en el interior de una campana de cristal y se introducen vapores químicos. El susceptor se hace girar y se aplica calor hasta que la capa epitaxial tiene el espesor suficiente. Se corrió un experimento utilizando dos factores: rapidez de flujo de arsénico (A) y tiempo de deposición (B). Se corrieron cuatro réplicas y se midió el espesor de la capa epitaxial (en μm). Los datos se muestran a continuación:

		Réplica					Niveles de factores	
A	B	I	II	III	IV		Bajo (–)	Alto (+)
–	–	14.037	16.165	13.972	13.907	A	55%	59%
+	–	13.880	13.860	14.032	13.914			
–	+	14.821	14.757	14.843	14.878	B	Corto	Largo
+	+	14.888	14.921	14.415	14.932		(10 min)	(15 min)

a) Estimar los efectos de los factores.
b) Conducir un análisis de varianza. ¿Qué factores son importantes?
c) Escribir una ecuación de regresión que podría usarse para predecir el espesor de la capa epitaxial en la región de la velocidad de flujo del arsénico y el tiempo de deposición utilizado en este experimento.
d) Analizar los residuales. ¿Se observa algún residual que debiera causar preocupación?
e) Comentar la forma en que se podría resolver el punto atípico potencial encontrado en el inciso *d*.

6-13. ***Continuación del problema 6-12.*** Usar el modelo de regresión del inciso *c* del problema 6-12 para generar una gráfica de contorno de la superficie de respuesta para el espesor epitaxial. Suponga que es de importancia crítica obtener un espesor de la capa de 14.5 μm. ¿Qué ajustes de la velocidad de flujo del arsénico y del tiempo de deposición se recomendarían?

6-14. ***Continuación del problema 6-13.*** ¿En qué forma cambiaría la respuesta dada en el problema 6-13 si la velocidad de flujo del arsénico fuera más difícil de controlar en el proceso que el tiempo de deposición?

6-15. Se utiliza una aleación de níquel y titanio para fabricar componentes de los motores de turbina de aviones. La formación de fisuras es un problema potencialmente serio de las piezas terminadas, ya que pueden provocar fallas irreversibles. Se realiza una prueba de las piezas para determinar el efecto de cuatro factores sobre las fisuras. Los cuatro factores son la temperatura de vaciado (A), el contenido de titanio (B), el método de tratamiento térmico (C) y la cantidad de refinador de grano usada (D). Se hacen dos réplicas de un diseño 2^4 y se mide la longitud de las fisuras (en mm $\times 10^{-2}$) inducidas en un ejemplar de prueba de muestra sometido a una prueba estándar. Los datos se muestran en la siguiente tabla:

A	B	C	D	Combinación de tratamientos	Réplica I	Réplica II
−	−	−	−	(1)	7.037	6.376
+	−	−	−	*a*	14.707	15.219
−	+	−	−	*b*	11.635	12.089
+	+	−	−	*ab*	17.273	17.815
−	−	+	−	*c*	10.403	10.151
+	−	+	−	*ac*	4.368	4.098
−	+	+	−	*bc*	9.360	9.253
+	+	+	−	*abc*	13.440	12.923
−	−	−	+	*d*	8.561	8.951
+	−	−	+	*ad*	16.867	17.052
−	+	−	+	*bd*	13.876	13.658
+	+	−	+	*abd*	19.824	19.639
−	−	+	+	*cd*	11.846	12.337
+	−	+	+	*acd*	6.125	5.904
−	+	+	+	*bcd*	11.190	10.935
+	+	+	+	*abcd*	15.653	15.053

a) Estimar los efectos de los factores. ¿Qué efectos de los factores parecen ser grandes?
b) Conducir un análisis de varianza. ¿Alguno de los factores afecta la formación de fisuras? Utilizar $\alpha = 0.05$.
c) Escribir un modelo de regresión que pueda usarse para predecir la longitud de las fisuras como una función de los efectos principales y las interacciones significativas que se han identificado en el inciso *b*.
d) Analizar los residuales de este experimento.
e) ¿Hay algún indicio de que alguno de los factores afecte la variabilidad de la formación de fisuras?
f) ¿Qué recomendaciones se harían respecto de las operaciones del proceso? Utilizar gráficas de las interacciones y/o de los efectos principales como ayuda para sacar conclusiones.

6-16. ***Continuación del problema 6-15.*** Una de las variables del experimento descrito en el problema 6-15, el método de tratamiento térmico (C), es una variable categórica. Suponga que los demás factores son continuos.
a) Escribir dos modelos de regresión para predecir la longitud de las fisuras, uno para cada nivel de la variable método de tratamiento térmico. ¿Qué diferencias, en caso de haberlas, se observan en estas dos ecuaciones?
b) Generar las gráficas de contorno apropiadas de la superficie de respuesta para los dos modelos de regresión del inciso *a*.
c) ¿Qué conjunto de condiciones se recomendaría para los factores A, B y D si se utiliza el método de tratamiento térmico $C = +$?
d) Repetir el inciso *c* suponiendo que quiere usarse el método de tratamiento térmico $C = -$.

6-17. Un experimentador corre una sola réplica de un diseño 2^4. Se calcularon las siguientes estimaciones de los efectos:

$A = 76.95$	$AB = -51.32$	$ABC = -2.82$
$B = -67.52$	$AC = 11.69$	$ABD = -6.50$
$C = -7.84$	$AD = 9.78$	$ACD = 10.20$
$D = -18.73$	$BC = 20.78$	$BCD = -7.98$
	$BD = 14.74$	$ABCD = -6.25$
	$CD = 1.27$	

a) Construir una gráfica de probabilidad normal de estos efectos.
b) Identificar un modelo tentativo, con base en la gráfica de los efectos del inciso *a*.

6-18. En un artículo de *Solid State Technology* ("Diseño ortogonal para optimización de procesos y su aplicación en el grabado químico con plasma") se describe la aplicación de diseños factoriales en el desarrollo de un proceso de grabado químico con nitruros en un dispositivo de grabado químico con plasma para una sola oblea. El proceso usa C_2F_6 como gas de reacción. Cuatro factores son de interés: el entrehierro ánodo-cátodo (A), la presión en la cámara del reactor (B), el flujo del gas C_2F_6 (C) y la potencia aplicada al cátodo (D). La respuesta de interés es la rapidez de grabado para el nitruro de silicio. Se corre una sola réplica de un diseño 2^4; los datos se muestran enseguida:

Número de corrida	Orden real de la corrida	A	B	C	D	Rapidez de grabado (Å/min)	Niveles de los factores	Bajo (−)	Alto (+)
1	13	−	−	−	−	550	A (cm)	0.80	1.20
2	8	+	−	−	−	669	B (mTorr)	450	550
3	12	−	+	−	−	604	C (SCCM)	125	200
4	9	+	+	−	−	650	D (W)	275	325
5	4	−	−	+	−	633			
6	15	+	−	+	−	642			
7	16	−	+	+	−	601			
8	3	+	+	+	−	635			
9	1	−	−	−	+	1037			
10	14	+	−	−	+	749			
11	5	−	+	−	+	1052			
12	10	+	+	−	+	868			
13	11	−	−	+	+	1075			
14	2	+	−	+	+	860			
15	7	−	+	+	+	1063			
16	6	+	+	+	+	729			

a) Estimar los efectos de los factores. Considere una gráfica de probabilidad normal de los efectos de los factores. ¿Qué efectos parecen ser grandes?
b) Efectuar un análisis de varianza para confirmar los resultados obtenidos en el inciso *a*.
c) ¿Cuál es el modelo de regresión que relaciona la rapidez de grabado con las variables significativas del proceso?
d) Analizar los residuales de este experimento. Comentar la adecuación del modelo.
e) Si no todos los factores son importantes, hacer la proyección del diseño 2^4 en un diseño 2^k con $k < 4$ y conducir el análisis de varianza.
f) Trazar gráficas para interpretar cualquier interacción significativa.
g) Graficar los residuales contra el orden real de las corridas. ¿Qué problemas podrían ponerse de manifiesto en esta gráfica?

6-19. ***Continuación del problema 6-18.*** Considere el modelo de regresión obtenido en el inciso *c* del problema 6-18.
a) Construir las gráficas de contorno de la rapidez de grabado utilizando este modelo.
b) Suponga que fuera necesario operar este proceso con una rapidez de 800 Å/min. ¿Cuáles serían los ajustes de las variables del proceso que se recomendarían?

6-20. Considere la réplica única del diseño 2^4 del ejemplo 6-2. Suponga que se decidió arbitrariamente analizar los datos suponiendo que las interacciones de tres y cuatro factores eran insignificantes. Conducir este análisis y comparar los resultados con los que se obtuvieron en el ejemplo. ¿Piensa el lector que es una buena idea suponer de manera arbitraria que las interacciones son insignificantes incluso cuando sean de orden relativamente alto?

6-21. Se realizó un experimento en una fábrica de semiconductores en un esfuerzo para incrementar el rendimiento. Se estudiaron cinco factores, cada uno con dos niveles. Los factores (y los niveles) fueron: A = ajuste de apertura (pequeña, grande), B = tiempo de exposición (20% abajo del nominal, 20% arriba del nominal),

C = tiempo de desarrollo (30 s, 45 s), D = tamaño de la máscara (pequeña, grande) y E = tiempo de grabado (14.5 min, 15.5 min). Se corrió el diseño 2^5 no replicado que se muestra a continuación.

(1) = 7	d = 8	e = 8	de = 6
a = 9	ad = 10	ae = 12	ade = 10
b = 34	bd = 32	be = 35	bde = 30
ab = 55	abd = 50	abe = 52	$abde$ = 53
c = 16	cd = 18	ce = 15	cde = 15
ac = 20	acd = 21	ace = 22	$acde$ = 20
bc = 40	bcd = 44	bce = 45	$bcde$ = 41
abc = 60	$abcd$ = 61	$abce$ = 65	$abcde$ = 63

a) Construir una gráfica de probabilidad normal de las estimaciones de los efectos. ¿Qué efectos parecen ser grandes?
b) Efectuar un análisis de varianza para confirmar los resultados obtenidos en el inciso *a*.
c) Escribir el modelo de regresión que relacione el rendimiento con las variables significativas del proceso.
d) Graficar los residuales en papel probabilidad normal. ¿La gráfica es satisfactoria?
e) Graficar los residuales contra los rendimientos predichos y contra cada uno de los cinco factores. Comentar las gráficas.
f) Interpretar cualquier interacción significativa.
g) ¿Qué recomendaciones se harían respecto de las condiciones de operación del proceso?
h) Hacer la proyección del diseño 2^5 de este problema en un diseño 2^k en los factores importantes. Esquematizar el diseño e indicar el promedio y el rango de los rendimientos en cada corrida. ¿Es de ayuda este esquema para interpretar los resultados de este experimento?

6-22. ***Continuación del problema 6-21.*** Suponga que el experimentador corrió cuatro puntos centrales además de los 32 ensayos del experimento original. Los rendimientos obtenidos en las corridas de los puntos centrales fueron 68, 74, 76 y 70.
a) Analizar de nuevo el experimento, incluyendo una prueba para la curvatura cuadrática pura.
b) Comentar cuál sería el siguiente paso.

6-23. Se estudiaron cuatro factores, cada uno con dos niveles, en un estudio del rendimiento de un proceso: el tiempo (A), la concentración (B), la presión (C) y la temperatura (D). Se corrió una sola réplica de un diseño 2^4 y los datos obtenidos se muestran en la siguiente tabla:

Número de corrida	Orden real de la corrida	A	B	C	D	Rendimiento (lbs)	Niveles de los factores	Bajo (−)	Alto (+)
1	5	−	−	−	−	12	A (h)	2.5	3
2	9	+	−	−	−	18	B (%)	14	18
3	8	−	+	−	−	13	C (psi)	60	80
4	13	+	+	−	−	16	D (°C)	225	250
5	3	−	−	+	−	17			
6	7	+	−	+	−	15			
7	14	−	+	+	−	20			
8	1	+	+	+	−	15			
9	6	−	−	−	+	10			
10	11	+	−	−	+	25			
11	2	−	+	−	+	13			
12	15	+	+	−	+	24			
13	4	−	−	+	+	19			
14	16	+	−	+	+	21			
15	10	−	+	+	+	17			
16	12	+	+	+	+	23			

a) Construir una gráfica de probabilidad normal de las estimaciones de los efectos. ¿Qué factores parecen tener efectos grandes?
b) Efectuar un análisis de varianza utilizando la gráfica de probabilidad normal del inciso *a* como guía para formar el término del error. ¿A qué conclusiones se llega?
c) Escribir un modelo de regresión que relacione el rendimiento con las variables importantes del proceso.
d) Analizar los residuales de este experimento. ¿El análisis indica algún problema potencial?
e) ¿Es posible plegar este diseño a un diseño 2^3 con dos réplicas? De ser así, esquematizar el diseño con el promedio y el rango del rendimiento indicados en cada punto del cubo. Interpretar los resultados.

6-24. ***Continuación del problema 6-23.*** Usar el modelo de regresión del inciso *c* del problema 6-23 para generar una gráfica de contorno de la superficie de respuesta del rendimiento. Analizar el valor práctico de esta gráfica de superficie de respuesta.

6-25. ***El experimento del brownie (pastelito) exquisito.*** El autor es un ingeniero hecho en la práctica y un firme creyente de aprender haciendo las cosas. Durante muchos años ha impartido el curso de diseño experimental a una amplia variedad de audiencias y siempre asigna la planeación, realización y análisis de un experimento real a los participantes de la clase. Los participantes parecen disfrutar esta experiencia práctica y siempre aprenden mucho de ella. En este problema se utilizan los resultados de un experimento realizado por Gretchen Krueger en la Universidad Estatal de Arizona.

Existen muchas formas diferentes de hornear brownies. El propósito de este experimento fue determinar la forma en que el material del molde, la marca de la harina para brownies y el método de batido afectan la exquisitez de los brownies. Los niveles de los factores fueron:

Factor	Bajo (–)	Alto (+)
A = material del molde	Vidrio	Aluminio
B = método de batido	Cuchara	Batidora
C = marca de la harina	Cara	Barata

La variable de respuesta fue la exquisitez, una medida subjetiva derivada de un cuestionario aplicado a los sujetos que hicieron el muestreo de cada lote de brownies. (Este cuestionario incluía aspectos como el sabor, la apariencia, la consistencia, el aroma, etc.) Un panel de prueba integrado por ocho personas hizo el muestreo de cada lote y llenó el cuestionario. La matriz del diseño y los datos de la respuesta se presentan a continuación:

Lote de brownies	A	B	C	Resultados del panel de prueba 1	2	3	4	5	6	7	8
1	–	–	–	11	9	10	10	11	10	8	9
2	+	–	–	15	10	16	14	12	9	6	15
3	–	+	–	9	12	11	11	11	11	11	12
4	+	+	–	16	17	15	12	13	13	11	11
5	–	–	+	10	11	15	8	6	8	9	14
6	+	–	+	12	13	14	13	9	13	14	9
7	–	+	+	10	12	13	10	7	7	17	13
8	+	+	+	15	12	15	13	12	12	9	14

a) Analizar los datos de este experimento como si se tratara de ocho réplicas de un diseño 2^3. Comentar los resultados.

b) ¿El análisis del inciso *a* es el enfoque correcto? Hay únicamente ocho lotes; ¿se tienen en realidad ocho réplicas de un diseño factorial 2^3?

c) Analizar el promedio y la desviación estándar del puntaje de la exquisitez. Comentar los resultados. ¿Este análisis es más apropiado que el del inciso *a*? ¿Por qué sí o no?

6-26. Se condujo un experimento en un proceso químico para producir un polímero. Los cuatro factores estudiados fueron la temperatura (*A*), la concentración del catalizador (*B*), el tiempo (*C*) y la presión (*D*). Se observaron dos respuestas, el peso molecular y la viscosidad. La matriz del diseño y los datos de la respuesta se presentan a continuación:

Número de corrida	Orden real de las corridas	*A*	*B*	*C*	*D*	Peso molecular	Viscosidad	Niveles de los factores	Bajo (–)	Alto (+)
1	18	–	–	–	–	2400	1400	*A* (°C)	100	120
2	9	+	–	–	–	2410	1500	*B* (%)	4	8
3	13	–	+	–	–	2315	1520	*C* (min)	20	30
4	8	+	+	–	–	2510	1630	*D* (psi)	60	75
5	3	–	–	+	–	2615	1380			
6	11	+	–	+	–	2625	1525			
7	14	–	+	+	–	2400	1500			
8	17	+	+	+	–	2750	1620			
9	6	–	–	–	+	2400	1400			
10	7	+	–	–	+	2390	1525			
11	2	–	+	–	+	2300	1500			
12	10	+	+	–	+	2520	1500			
13	4	–	–	+	+	2625	1420			
14	19	+	–	+	+	2630	1490			
15	15	–	+	+	+	2500	1500			
16	20	+	+	+	+	2710	1600			
17	1	0	0	0	0	2515	1500			
18	5	0	0	0	0	2500	1460			
19	16	0	0	0	0	2400	1525			
20	12	0	0	0	0	2475	1500			

a) Considere únicamente la respuesta del peso molecular. Graficar las estimaciones de los efectos en una escala de probabilidad normal. ¿Qué efectos parecen ser importantes?

b) Usar un análisis de varianza para confirmar los resultados del inciso *a*. ¿Hay algún indicio de curvatura?

c) Escribir un modelo de regresión para predecir el peso molecular como una función de las variables importantes.

d) Analizar los residuos y comentar la adecuación del modelo.

e) Repetir los incisos *a–d* utilizando la respuesta de la viscosidad.

6-27. ***Continuación del problema 6-26.*** Utilizar los modelos de regresión del peso molecular y la viscosidad para responder las preguntas siguientes.

a) Construir una gráfica de contorno de la superficie de respuesta para el peso molecular. ¿En qué dirección se ajustarían las variables del proceso a fin de incrementar el peso molecular?

b) Construir una gráfica de contorno de la superficie de respuesta para la viscosidad. ¿En qué dirección se ajustarían las variables del proceso para disminuir la viscosidad?

c) ¿Qué condiciones de operación se recomendarían si fuera necesario producir un producto con peso molecular entre 2400 y 2500, y con la viscosidad más baja posible?

6-28. Considere una sola réplica del diseño 2^4 del ejemplo 6-2. Suponga que se hicieron cinco corridas de puntos en el centro (0, 0, 0, 0) y que se observaron las respuestas siguientes: 73, 75, 71, 69 y 76. Probar la curvatura en este experimento. Interpretar los resultados.

6-29. ***Un valor faltante en un diseño factorial 2^k.*** No es raro encontrar que falta una de las observaciones de un diseño 2^k debido a un equipo de medición defectuoso, una prueba fallida, o alguna otra razón. Si el diseño se hace con n réplicas ($n > 1$), puede emplearse alguna de las técnicas estudiadas en el capítulo 5. Sin embargo, para un diseño factorial sin réplicas ($n = 1$) debe usarse otro método. Un enfoque lógico es estimar el valor faltante con un número que haga cero el contraste de la interacción de orden más alto. Aplicar esta técnica al experimento del ejemplo 6-2, suponiendo que falta la corrida ab. Compare los resultados obtenidos con los del ejemplo 6-2.

6-30. Un ingeniero realizó un experimento para estudiar el efecto de cuatro factores sobre la aspereza superficial de una pieza maquinada. Los factores (y sus niveles) son A = ángulo de la herramienta (12, 15°), B = viscosidad del fluido de corte (300, 400), C = velocidad de alimentación (10, 15 pulg/min) y D = enfriador del fluido de corte usado (no, sí). Los datos de este experimento (con los factores codificados en los niveles usuales −1, +1) se muestran a continuación.

Corrida	A	B	C	D	Rugosidad superficial
1	−	−	−	−	0.00340
2	+	−	−	−	0.00362
3	−	+	−	−	0.00301
4	+	+	−	−	0.00182
5	−	−	+	−	0.00280
6	+	−	+	−	0.00290
7	−	+	+	−	0.00252
8	+	+	+	−	0.00160
9	−	−	−	+	0.00336
10	+	−	−	+	0.00344
11	−	+	−	+	0.00308
12	+	+	−	+	0.00184
13	−	−	+	+	0.00269
14	+	−	+	+	0.00284
15	−	+	+	+	0.00253
16	+	+	+	+	0.00163

a) Estimar los efectos de los factores. Representar las efectos de los factores en una gráfica de probabilidad normal y seleccionar un modelo tentativo.

b) Ajustar el modelo identificado en el inciso a y analizar los residuales. ¿Hay algún indicio de que el modelo no sea adecuado?

c) Repetir el análisis de los incisos a y b utilizando $1/y$ como la variable de respuesta. ¿Hay algún indicio de que la transformación ha sido útil?

d) Ajustar un modelo en términos de las variables codificadas que pueda usarse para predecir la rugosidad superficial. Convertir esta ecuación de predicción en un modelo en las variables naturales.

6-31. La resistividad de una oblea de silicio está influida por varios factores. Los resultados de un experimento factorial 2^4 realizado durante un paso crítico del procesamiento se muestran en la tabla siguiente:

Corrida	A	B	C	D	Resistividad
1	−	−	−	−	1.92
2	+	−	−	−	11.28
3	−	+	−	−	1.09
4	+	+	−	−	5.75
5	−	−	+	−	2.13
6	+	−	+	−	9.53
7	−	+	+	−	1.03
8	+	+	+	−	5.35
9	−	−	−	+	1.60
10	+	−	−	+	11.73
11	−	+	−	+	1.16
12	+	+	−	+	4.68
13	−	−	+	+	2.16
14	+	−	+	+	9.11
15	−	+	+	+	1.07
16	+	+	+	+	5.30

a) Estimar los efectos de los factores. Representar las efectos de los factores en una gráfica de probabilidad normal y seleccionar un modelo tentativo.
b) Ajustar el modelo identificado en el inciso *a* y analizar los residuales. ¿Hay algún indicio de que el modelo no sea adecuado?
c) Repetir el análisis de los incisos *a* y *b* utilizando ln (y) como la variable de respuesta. ¿Hay algún indicio de que la transformación haya sido útil?
d) Ajustar un modelo en términos de las variables codificadas que pueda usarse para predecir la resistividad.

6-32. ***Continuación del problema 6-31.*** Suponga que el experimentador corrió también cuatro puntos centrales junto con las 16 corridas del problema 6-31. Las mediciones de la resistividad en los puntos centrales son: 8.15, 7.63, 8.95 y 6.48. Analizar de nuevo el experimento incorporando los puntos centrales. ¿Qué conclusiones pueden sacarse ahora?

6-33. Es frecuente usar el modelo de regresión ajustado de un diseño factorial 2^k para hacer predicciones en puntos de interés del espacio del diseño.
a) Encontrar la varianza de la respuesta predicha $\hat{y}$ en un punto $x_1, x_2, ..., x_k$ del espacio del diseño. *Sugerencia*: recuerde que las x están codificadas, y suponga un diseño 2^k con el mismo número de réplicas n en cada punto del diseño, de tal modo que la varianza de un coeficiente de regresión $\hat{\beta}$ sea $\sigma^2/(n2^k)$ y que la covarianza entre cualquier par de coeficientes de regresión sea cero.
b) Usar el resultado del inciso *a* para encontrar la ecuación de un intervalo de confianza de $100(1-\alpha)$ por ciento para la verdadera respuesta media en el punto $x_1, x_2, ..., x_k$ del espacio del diseño.

6-34. ***Modelos jerárquicos.*** Se ha usado varias veces el principio de jerarquía para seleccionar un modelo; es decir, se han incluido términos de orden inferior no significativos en un modelo porque eran factores que estaban incluidos en términos de orden superior significativos. Ciertamente, la jerarquía no es un principio absoluto que deba seguirse en todos los casos. Para ilustrar esto, considere el modelo que resultó en el problema 6-1, el cual requirió que se incluyera un efecto principal no significativo para respetar la jerarquía. Utilizar los datos del problema 6-1.
a) Ajustar el modelo jerárquico y el modelo no jerárquico.
b) Calcular el estadístico PRESS, la R^2 ajustada y el cuadrado medio del error para ambos modelos.
c) Encontrar un intervalo de confianza de 95% para la estimación de la respuesta media en el vértice de un cubo ($x_1 = x_2 = x_3 = \pm 1$). *Sugerencia*: usar los resultados del problema 6-33.
d) Con base en los análisis que se han realizado, ¿qué modelo preferiría el lector?

7

Formación de bloques y confusión en el diseño factorial 2^k

7-1 INTRODUCCIÓN

Hay múltiples situaciones en las que es imposible efectuar todas las corridas de un experimento factorial 2^k bajo condiciones homogéneas. Por ejemplo, un lote de materia prima podría no ser suficiente para hacer todas las corridas requeridas. En otros casos, podría ser conveniente modificar deliberadamente las condiciones experimentales para asegurar que los tratamientos tengan la misma efectividad (es decir, que sean robustos) en diversas situaciones que es posible encontrar en la práctica. Por ejemplo, un ingeniero químico puede correr un experimento en una planta piloto con varios lotes de materia prima porque sabe que en el proceso real a gran escala posiblemente se usarán diferentes lotes de materia prima con diversos grados de calidad.

La técnica de diseño utilizada en estas situaciones es la **formación de bloques.** Este capítulo se enfoca en algunas técnicas especiales para separar en bloques un diseño factorial 2^k.

7-2 FORMACIÓN DE BLOQUES DE UN DISEÑO FACTORIAL 2^k CON RÉPLICAS

Suponga que se han corrido n réplicas del diseño factorial 2^k. Esta situación es idéntica a la que se estudió en el capítulo 5, donde se indicó cómo correr un diseño factorial general en bloques. Si hay n réplicas, entonces cada conjunto de condiciones no homogéneas define un bloque, y cada réplica se corre en uno de los bloques. Las corridas de cada bloque (o réplica) se harían de manera aleatoria. El análisis del diseño

Tabla 7-1 Experimento del proceso químico en tres bloques

	Bloque 1	Bloque 2	Bloque 3
	(1) = 28	(1) = 25	(1) = 27
	a = 36	a = 32	a = 32
	b = 18	b = 19	b = 23
	ab = 31	ab = 30	ab = 29
Totales de los bloques	B_1 = 113	B_2 = 106	B_3 = 111

Tabla 7-2 Análisis de varianza del experimento del proceso químico en tres bloques

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0	Valor P
Bloques	6.50	2	3.25		
A (concentración)	208.33	1	208.33	50.32	0.0004
B (catalizador)	75.00	1	75.00	18.12	0.0053
AB	8.33	1	8.33	2.01	0.2060
Error	24.84	6	4.14		
Total	323.00	11			

es similar al de cualquier experimento factorial separado en bloques; por ejemplo, véase la revisión de la sección 5-6.

EJEMPLO 7-1 ....

Considere el experimento del proceso químico que se describió en la sección 6-2. Suponga que sólo pueden hacerse cuatro ensayos experimentales con un solo lote de materia prima. Por lo tanto, se necesitarán tres lotes de materia prima para correr las tres réplicas de este diseño. En la tabla 7-1 se muestra el diseño donde cada lote de materia prima corresponde a un bloque.

En la tabla 7-2 se muestra el análisis de varianza de este diseño separado en bloques. Todas las sumas de cuadrados se calculan exactamente igual que en un diseño 2^k estándar sin formación de bloques. La suma de cuadrados de los bloques se calcula a partir de los totales de los bloques. Sea que B_1, B_2 y B_3 representen los totales de los bloques (ver la tabla 7-1). Entonces

$$
\begin{aligned}
SS_{\text{Bloques}} &= \sum_{i=1}^{3} \frac{B_i^2}{4} - \frac{y_{...}^2}{12} \\
&= \frac{(113)^2 + (106)^2 + (111)^2}{4} - \frac{(330)^2}{12} \\
&= 6.50
\end{aligned}
$$

Hay dos grados de libertad entre los tres bloques. La tabla 7-2 indica que las conclusiones de este análisis, si el diseño se hubiera corrido en bloques, son idénticas a las de la sección 6-2 y que el efecto de los bloques es relativamente pequeño.

....

7-3 CONFUSIÓN DEL DISEÑO FACTORIAL 2^k

Hay muchos problemas en los que es imposible realizar una réplica completa de un diseño factorial en un bloque. La **confusión** (o **mezclado**) es una técnica de diseño mediante la cual un experimento factorial completo se distribuye en bloques, donde el tamaño del bloque es menor que el número de combinaciones de los tratamientos de una réplica. La técnica hace que la información acerca de ciertos efectos de los tratamientos (por lo general las interacciones de orden superior) sea **indistinguible de los bloques** o esté **confundida con los bloques.** En este capítulo la atención se centra en los sistemas de confusión (o mezclado) para el diseño factorial 2^k. Observe que aun cuando los diseños que se presentan son diseños de bloques incompletos, ya que cada bloque no contiene todos los tratamientos o las combinaciones de los tratamientos, la estructura especial del sistema factorial 2^k permite un método de análisis simplificado.

Se considera la construcción y el análisis del diseño factorial 2^k en 2^p bloques incompletos, donde $p < k$. Por consiguiente, estos diseños pueden correrse en dos bloques, en cuatro bloques, en ocho bloques, etcétera.

7-4 CONFUSIÓN DEL DISEÑO FACTORIAL 2^k EN DOS BLOQUES

Suponga que quiere correrse una sola réplica del diseño 2^2. Cada una de las $2^2 = 4$ combinaciones de los tratamientos requiere una cantidad de materia prima, por ejemplo, y cada lote de materia prima sólo alcanza para probar dos combinaciones de tratamientos. Por lo tanto, se necesitan dos lotes de materia prima. Si los lotes de materia prima se consideran como bloques, entonces deben asignarse a cada bloque dos de las cuatro combinaciones de tratamientos.

En la figura 7-1 se muestra uno de los diseños posibles para este problema. La vista geométrica, figura 7-1*a*, indica que las combinaciones de tratamientos localizadas en diagonales opuestas se asignan a bloques diferentes. Observe, por la figura 7-1*b*, que el bloque 1 contiene las combinaciones de los tratamientos (1) y *ab* y que el bloque 2 contiene *a* y *b*. Desde luego, el *orden* en que se corren las combinaciones de los tratamientos dentro de un bloque se determina aleatoriamente. También se decidirá aleatoriamente cuál de los bloques se correrá primero. Suponga que los efectos principales de A y B se estiman como si no se hubiera hecho la formación de bloques. Por las ecuaciones 6-1 y 6-2 se obtiene

$$A = \tfrac{1}{2}[ab + a - b - (1)]$$
$$B = \tfrac{1}{2}[ab + b - a - (1)]$$

Observe que ni A ni B son afectados por la formación de bloques, debido a que en cada estimación hay una combinación de un tratamiento positivo y uno negativo de cada bloque. Es decir, cualquier diferencia entre el bloque 1 y el bloque 2 se cancela.

Considere ahora la interacción AB

$$AB = \tfrac{1}{2}[ab + (1) - a - b]$$

Puesto que las dos combinaciones de tratamientos con signo positivo [*ab* y (1)] están en el bloque 1 y las dos con signo negativo (*a* y *b*) están en el bloque 2, el efecto de los bloques y la interacción AB son idénticos. Es decir, AB está **confundido** (o **mezclado**) con los bloques.

La razón de esto es evidente en la tabla de signos positivos y negativos del diseño 2^2. Se presentó originalmente en la tabla 6-2, pero por conveniencia se repite como la tabla 7-3. A partir de esta tabla se observa que todas las combinaciones de tratamientos que tienen signo positivo para AB se asignan al bloque 1,

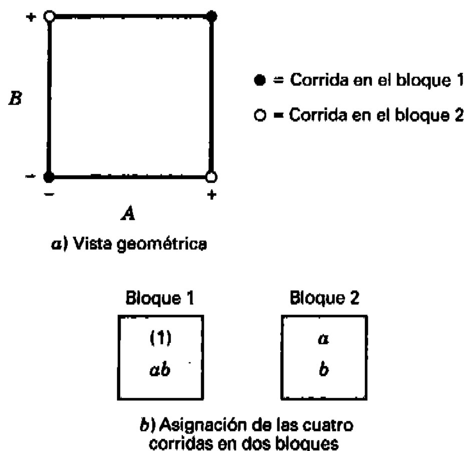


Figura 7-1 Diseño 2^2 en dos bloques.

Tabla 7-3 Tabla de signos positivos y negativos para el diseño 2^2

Combinación de tratamientos	Efecto factorial			
	I	*A*	*B*	*AB*
(1)	+	−	−	+
a	+	+	−	−
b	+	−	+	−
ab	+	+	+	+

mientras que todas las combinaciones de tratamientos que tienen signo negativo para *AB* se asignan al bloque 2. Este enfoque puede usarse para confundir o mezclar cualquier efecto (*A*, *B* o *AB*) con los bloques. Por ejemplo, si (1) y *b* se hubieran asignado al bloque 1 y *a* y *ab* al bloque 2, el efecto principal de *A* se habría confundido con los bloques. La práctica usual es confundir la interacción de orden más alto con los bloques.

Este esquema puede usarse para confundir o mezclar cualquier diseño 2^k en dos bloques. Como un segundo ejemplo, considere un diseño 2^3 que se corre en dos bloques. Suponga que se quiere confundir la interacción de los tres factores *ABC* con los bloques. Por la formación de signos positivos y negativos de la tabla 7-4, las combinaciones de tratamientos que son negativas para *ABC* se asignan al bloque 1 y las que son positivas para *ABC* al bloque 2. El diseño resultante se muestra en la figura 7-2. De nueva cuenta se resalta que las combinaciones de tratamientos *dentro* de un bloque se corren de manera aleatoria.

Otros métodos para construir bloques

Se cuenta con otro método para construir estos diseños. El método utiliza la combinación lineal

$$L = \alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_k x_k \tag{7-1}$$

donde x_i es el nivel del factor *i*-ésimo que aparece en una combinación de tratamientos particular y α_i es el exponente que aparece en el factor *i*-ésimo para el efecto que va a confundirse. Para el sistema 2^k, se tiene $\alpha_i = 0$ o 1 y $x_i = 0$ (nivel bajo) o $x_i = 1$ (nivel alto). A la ecuación 7-1 se le llama la **definición de contrastes.** Las combinaciones de tratamientos que producen el mismo valor de L (mod 2) se colocarán en el mismo bloque. Puesto que los únicos valores posibles de L (mod 2) son 0 y 1, con esto las 2^k combinaciones de tratamientos se asignarán a exactamente dos bloques.

Tabla 7-4 Tabla de signos positivos y negativos para el diseño 2^3

Combinación de tratamientos	Efecto factorial							
	I	*A*	*B*	*AB*	*C*	*AC*	*BC*	*ABC*
(1)	+	−	−	+	−	+	+	−
a	+	+	−	−	−	−	+	+
b	+	−	+	−	−	+	−	+
ab	+	+	+	+	−	−	−	−
c	+	−	−	+	+	−	−	+
ac	+	+	−	−	+	+	−	−
bc	+	−	+	−	+	−	+	−
abc	+	+	+	+	+	+	+	+

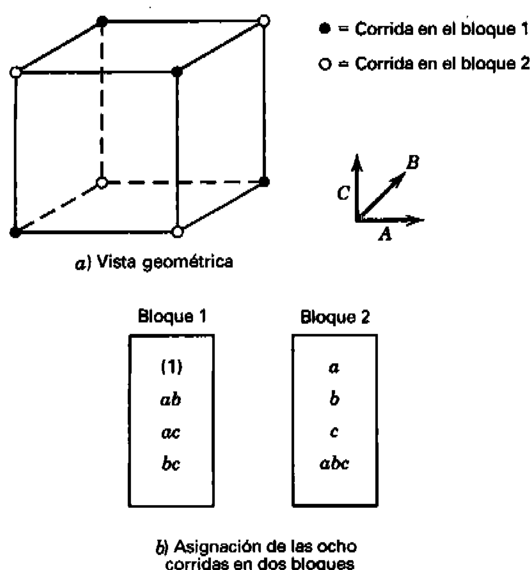


Figura 7-2 El diseño 2^3 en dos bloques con ABC confundido.

Para ilustrar este enfoque, considere un diseño 2^3 con ABC confundido con los bloques. En este caso, x_1 corresponde a A, x_2 a B, x_3 a C y $\alpha_1 = \alpha_2 = \alpha_3 = 1$. Por lo tanto, la definición del contraste correspondiente a ABC es

$$L = x_1 + x_2 + x_3$$

La combinación de tratamientos (1) se escribe 000 en la notación (0, 1); por lo tanto,

$$L = 1(0) + 1(0) + 1(0) = 0 = 0 \text{ (mod 2)}$$

De manera similar, la combinación de tratamientos a es 100, obteniéndose

$$L = 1(1) + 1(0) + 1(0) = 1 = 1 \text{ (mod 2)}$$

Por lo tanto, (1) y a se correrían en bloques diferentes. Para el resto de las combinaciones de tratamientos se tiene

$$\begin{aligned}
b&: \quad L = 1(0) + 1(1) + 1(0) = 1 = 1 \text{ (mod 2)} \\
ab&: \quad L = 1(1) + 1(1) + 1(0) = 2 = 0 \text{ (mod 2)} \\
c&: \quad L = 1(0) + 1(0) + 1(1) = 1 = 1 \text{ (mod 2)} \\
ac&: \quad L = 1(1) + 1(0) + 1(1) = 2 = 0 \text{ (mod 2)} \\
bc&: \quad L = 1(0) + 1(1) + 1(1) = 2 = 0 \text{ (mod 2)} \\
abc&: \quad L = 1(1) + 1(1) + 1(1) = 3 = 1 \text{ (mod 2)}
\end{aligned}$$

Por lo tanto, (1), ab, ac y bc se corren en el bloque 1 y a, b, c y abc se corren en el bloque 2. Se trata del mismo diseño que se ilustró en la figura 7-2, el cual se generó con la tabla de signos positivos y negativos.

Puede usarse otro método para construir estos diseños. Al bloque que contiene la combinación de tratamientos (1) se le llama el **bloque principal.** Las combinaciones de los tratamientos incluidas en este bloque poseen una útil propiedad de la teoría de grupos; a saber, forman un grupo con respecto a la multi-

plicación módulo 2. Esto implica que cualquier elemento [con excepción de (1)] del bloque principal puede generarse multiplicando otros dos elementos del bloque principal módulo 2. Por ejemplo, considere el bloque principal del diseño 2^3 con *ABC* confundido, como se muestra en la figura 7-2. Observe que

$$ab \cdot ac = a^2bc = bc$$
$$ab \cdot bc = ab^2c = ac$$
$$ac \cdot bc = abc^2 = ab$$

Las combinaciones de tratamientos del otro bloque (o bloques) pueden generarse multiplicando uno de los elementos del nuevo bloque por cada uno de los elementos del bloque principal módulo 2. Para el diseño 2^3 con *ABC* confundido, puesto que el bloque principal es (1), *ab, ac* y *bc*, se sabe que *b* está en el otro bloque. Por lo tanto, los elementos de este segundo bloque son

$$b \cdot (1) \quad = b$$
$$b \cdot ab = ab^2 = a$$
$$b \cdot ac \quad = abc$$
$$b \cdot bc = b^2c = c$$

Estos resultados concuerdan con los que se obtuvieron anteriormente.

Estimación del error

Cuando el número de variables es pequeño, por ejemplo $k = 2$ o 3, por lo general es necesario hacer réplicas del experimento a fin de obtener una estimación del error. Por ejemplo, suponga que un diseño factorial 2^3 debe correrse en dos bloques con *ABC* confundido, y el experimentador decide hacer cuatro réplicas del diseño. El diseño resultante podría verse como el de la figura 7-3. Observe que *ABC* está confundido en cada réplica.

En la tabla 7-5 se muestra el análisis de varianza de este diseño. Hay 32 observaciones y 31 grados de libertad. Además, puesto que hay ocho bloques, siete grados de libertad deben asociarse con estos bloques. En la tabla 7-5 se presenta la descomposición de esos siete grados de libertad. La suma de cuadrados del error se compone en realidad de las interacciones de dos factores entre las réplicas, y cada uno de los efectos (*A, B, C, AB, AC, BC*). Por lo general es seguro considerar que las interacciones son cero y tratar el cuadrado medio resultante como una estimación del error. Los efectos principales y las interacciones de dos factores se prueban contra el cuadrado medio del error. Cochran y Cox [25b] hacen notar que el cuadrado medio del bloque o *ABC* podría compararse con el error del cuadrado medio *ABC*, que es en realidad réplicas × bloques. Esta prueba suele tener una sensibilidad muy baja.

Si se cuenta con recursos suficientes para hacer réplicas de un diseño confundido, por lo general es mejor usar un método ligeramente diferente para diseñar los bloques en cada réplica. Este enfoque consiste en confundir un efecto diferente en cada réplica para obtener cierta información sobre todos los

Réplica I		Réplica II		Réplica III		Réplica IV	
Bloque 1	Bloque 2	Bloque 1	Bloque 2	Bloque 1	Bloque 2	Bloque 1	Bloque 2
(1)	*abc*	(1)	*abc*	(1)	*abc*	(1)	*abc*
ac	*a*	*ac*	*a*	*ac*	*a*	*ac*	*a*
ab	*b*	*ab*	*b*	*ab*	*b*	*ab*	*b*
bc	*c*	*bc*	*c*	*bc*	*c*	*bc*	*c*

Figura 7-3 Cuatro réplicas del diseño 2^3 con *ABC* confundido.

Tabla 7-5 Análisis de varianza de cuatro réplicas de un diseño 2^3 con *ABC* confundido

Fuente de variación	Grados de libertad
Réplicas	3
Bloques (*ABC*)	1
Error de *ABC* (réplicas × bloques)	3
A	1
B	1
C	1
AB	1
AC	1
BC	1
Error (o réplicas × efectos)	18
Total	31

efectos. A este procedimiento se le llama **confusión** (o **mezclado**) **parcial**, y se estudia en la sección 7-7. Si k es moderadamente grande, por ejemplo $k \geq 4$, con frecuencia sólo es posible hacer una réplica. El experimentador suele suponer que las interacciones de órdenes superiores son insignificantes y combina sus sumas de cuadrados como el error. La gráfica de probabilidad normal de los efectos de los factores puede ser muy útil a este respecto.

EJEMPLO 7-2 ··

Considere la situación descrita en el ejemplo 6-2. Recuerde que se estudian cuatro factores—la temperatura (*A*), la presión (*B*), la concentración de formaldehído (*C*) y la velocidad de agitación (*D*)— en una planta piloto para determinar su efecto sobre el índice de filtración del producto. Se usará este experimento para ilustrar las ideas de la formación de bloques y la confusión en un diseño no replicado. Se introducirán dos modificaciones al experimento original. Primera, suponga que no es posible correr las $2^4 = 16$ combinaciones de tratamientos utilizando un solo lote de materia prima. El experimentador puede correr ocho combinaciones de los tratamientos con un solo lote de material, por lo que un diseño 2^4 confundido en dos bloques parece apropiado. Es lógico confundir la interacción de orden más alto *ABCD* con los bloques. La definición del contraste es

$$L = x_1 + x_2 + x_3 + x_4$$

y es sencillo verificar que el diseño es como el que se ilustra en la figura 7-4. De manera alternativa, puede examinarse la tabla 6-12 y observar que las combinaciones de los tratamientos que son + en la columna *ABCD* se asignan al bloque 1 y que las que son – en la columna *ABCD*, están en el bloque 2.

La segunda modificación que se hará es introducir un efecto de los bloques para que pueda demostrarse la utilidad de la formación de bloques. Suponga que cuando se seleccionan los dos lotes de materia prima que se necesitan para correr el experimento, uno de ellos es de calidad mucho más baja y, como resultado, todas las respuestas serán 20 unidades menores en este lote de material que en el otro. El lote de calidad menor se convierte en el bloque 1 y el lote de buena calidad se convierte en el bloque 2 (no es relevante a cuál de los dos lotes se le llama bloque 1 o bloque 2). Entonces todas las pruebas del bloque 1 se realizan primero (las ocho corridas del bloque se hacen, desde luego, de manera aleatoria), pero las respuestas son 20 unidades más bajas que las que se habrían obtenido si se hubiera usado el material de buena calidad. En la figura 7-4*b* se muestran las respuestas resultantes; observe que éstas se han encontrado

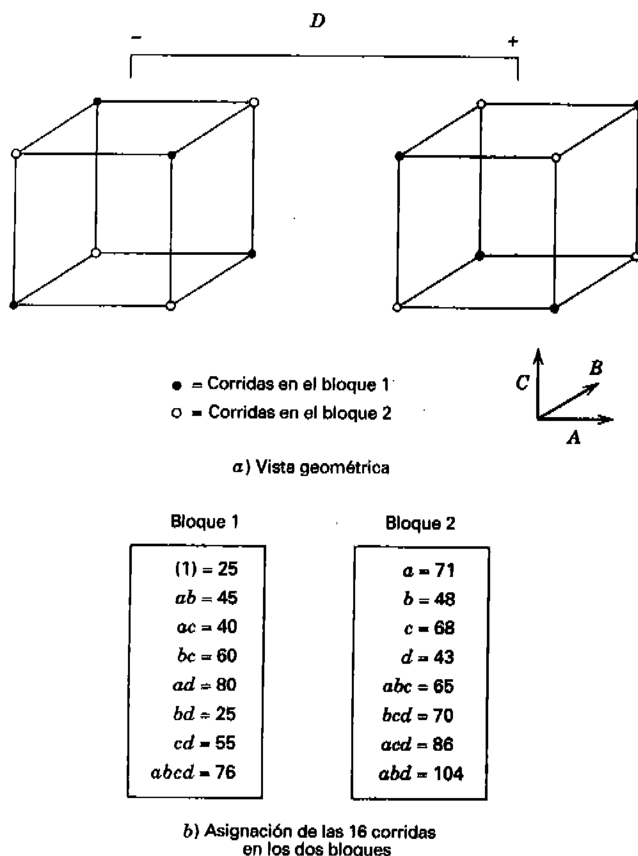


Figura 7-4 El diseño 2^4 en dos bloques para el ejemplo 7-2.

restando el efecto del bloque de las observaciones originales dadas en el ejemplo 6-2. Es decir, la respuesta original de la combinación de tratamientos (1) fue 45, y en la figura 7-4*b* se consigna como (1) = 25 (= 45 – 20). Las demás respuestas de este bloque se obtienen de manera similar. Después de que se realizan las pruebas del bloque 1, se prosigue con las ocho pruebas del bloque 2. No hay ningún problema con la materia prima de este lote, por lo que las respuestas son exactamente como fueron originalmente en el ejemplo 6-2.

En la tabla 7-6 se muestran las estimaciones de los efectos para esta versión "modificada" del ejemplo 6-2. Observe que las estimaciones de los cuatro efectos principales, de las seis interacciones de dos factores y de las cuatro interacciones de tres factores son idénticas a las estimaciones de los efectos obtenidas en el ejemplo 6-2, donde no hubo *ningún efecto de bloques*. Cuando se construye una gráfica de probabilidad normal de estas estimaciones de los efectos, los factores A, C, D y las interacciones AC y AD aparecen como los efectos importantes, justo como en el experimento original. (El lector deberá verificar esto.)

¿Qué puede decirse del efecto de la interacción $ABCD$? La estimación de este efecto en el experimento original (ejemplo 6-2) fue $ABCD = 1.375$. En el presente ejemplo, la estimación del efecto de la interacción $ABCD$ es $ABCD = -18.625$. Puesto que $ABCD$ está confundido con los bloques, la interacción $ABCD$ estima el *efecto de la interacción original* (1.375) más el *efecto de bloque* (–20), de donde $ABCD = 1.375 + (-20) = -18.625$. (¿Puede el lector ver por qué el efecto del bloque es –20?) El efecto del bloque

Tabla 7-6 Estimaciones de los efectos para el diseño 2^4 separado en bloques del ejemplo 7-2

Término del modelo	Coeficiente de regresión	Estimación del efecto	Suma de cuadrados	Contribución porcentual
A	10.81	21.625	1870.5625	26.30
B	1.56	3.125	39.0625	0.55
C	4.94	9.875	390.0625	5.49
D	7.31	14.625	855.5625	12.03
AB	0.062	0.125	0.0625	<0.01
AC	−9.06	−18.125	1314.0625	18.48
AD	8.31	16.625	1105.5625	15.55
BC	1.19	2.375	22.5625	0.32
BD	−0.19	−0.375	0.5625	<0.01
CD	−0.56	−1.125	5.0625	0.07
ABC	0.94	1.875	14.0625	0.20
ABD	2.06	4.125	68.0625	0.96
ACD	−0.81	−1.625	10.5625	0.15
BCD	−1.31	−2.625	27.5625	0.39
Bloques (*ABCD*)		−18.625	1387.5625	19.51

también puede calcularse directamente como la diferencia en la respuesta promedio entre los dos bloques, o

$$\begin{aligned}\text{Efecto del bloque} &= \bar{y}_{\text{Bloque 1}} - \bar{y}_{\text{Bloque 2}}\\ &= \frac{406}{8} - \frac{555}{8}\\ &= \frac{-149}{8}\\ &= -18.625\end{aligned}$$

Desde luego, este efecto es en realidad la estimación de Bloques + *ABCD*.

En la tabla 7-7 se resume el análisis de varianza de este experimento. Los efectos que tienen estimaciones grandes están incluidos en el modelo, y la suma de cuadrados de los bloques es

$$SS_{\text{Bloques}} = \frac{(406)^2 + (555)^2}{8} - \frac{(961)^2}{16} = 1387.5625$$

Tabla 7-7 Análisis de varianza del ejemplo 7-2

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0	Valor *P*
Bloques (*ABCD*)	1387.5625	1			
A	1870.5625	1	1870.5625	89.76	<0.0001
C	390.0625	1	390.0625	18.72	0.0019
D	855.5625	1	855.5625	41.05	0.0001
AC	1314.0625	1	1314.0625	63.05	<0.0001
AD	1105.5625	1	1105.5625	53.05	<0.0001
Error	187.5625	9	20.8403		
Total	7111.4375	15			

Las conclusiones de este experimento coinciden exactamente con las del ejemplo 6-2, donde no estuvo presente ningún efecto de bloques. Observe que si el experimento no se hubiera corrido en bloques, y si un efecto de magnitud –20 hubiera afectado los 8 primeros ensayos (los cuales se habrían seleccionado de manera aleatoria, ya que los 16 ensayos se habrían corrido en orden aleatorio en un diseño sin formación de bloques), los resultados pudieron haber sido muy diferentes.

7-5 CONFUSIÓN DEL DISEÑO FACTORIAL 2^k EN CUATRO BLOQUES

Es posible construir diseños factoriales 2^k confundidos en cuatro bloques con 2^{k-2} observaciones cada uno. Estos diseños son particularmente útiles en situaciones en las que el número de factores es moderadamente grande, por ejemplo $k \geq 4$, y el tamaño de los bloques es relativamente pequeño.

Como un ejemplo, considere el diseño 2^5. Si cada bloque incluirá únicamente ocho corridas, entonces deberán usarse cuatro bloques. La construcción de este diseño es relativamente directa. Se seleccionan *dos* efectos para confundirlos con los bloques, por ejemplo *ADE* y *BCE*. Estos efectos tienen las dos definiciones de contrastes

$$L_1 = x_1 + x_4 + x_5$$
$$L_2 = x_2 + x_3 + x_5$$

asociadas con ellos. Entonces cada combinación de tratamientos producirá un par particular de valores de L_1 (mod 2) y L_2 (mod 2), es decir, cualquiera de $(L_1, L_2) = (0, 0), (0, 1), (1, 0)$ o bien $(1, 1)$. Las combinaciones de tratamientos que producen los mismos valores de (L_1, L_2) se asignan al mismo bloque. En el ejemplo tratado aquí se encuentra

$L_1 = 0, L_2 = 0$ para (1), *ad, bc, abcd, abe, ace, cde, bde*
$L_1 = 1, L_2 = 0$ para *a, d, abc, bcd, be, abde, ce, acde*
$L_1 = 0, L_2 = 1$ para *b, abd, c, acd, ae, de, abce, bcde*
$L_1 = 1, L_2 = 1$ para *e, ade, bce, abcde, ab, bd, ac, cd*

Estas combinaciones de tratamientos se asignarían a bloques diferentes. En la figura 7-5 se muestra el diseño completo.

Con un poco de reflexión, nos damos cuenta de que otro efecto además de *ADE* y *BCE* debe confundirse con los bloques. Puesto que hay cuatro bloques con tres grados de libertad entre ellos, y puesto que *ADE* y *BCE* tienen un solo grado de libertad cada una, es evidente la necesidad de confundir un efecto adicional con un grado de libertad. Este efecto es la **interacción generalizada** de *ADE* y *BCE*, la cual se

Bloque 1		Bloque 2		Bloque 3		Bloque 4	
$L_1 = 0$		$L_1 = 1$		$L_1 = 0$		$L_1 = 1$	
$L_2 = 0$		$L_2 = 0$		$L_2 = 1$		$L_2 = 1$	
(1)	*abc*	*a*	*be*	*b*	*abce*	*e*	*abcde*
ad	*ace*	*d*	*abde*	*abd*	*ae*	*ade*	*bd*
bc	*cde*	*abc*	*ce*	*c*	*bcde*	*bce*	*ac*
abcd	*bde*	*bcd*	*acde*	*acd*	*de*	*ab*	*cd*

Figura 7-5 El diseño 2^5 en cuatro bloques con *ADE*, *BCE* y *ABCD* confundidos.

define como el producto de *ADE* y *BCE* módulo 2. Por lo tanto, en el ejemplo tratado aquí la interacción generalizada $(ADE)(BCE) = ABCDE^2 = ABCD$ también está confundido con los bloques. Es sencillo verificar esto refiriéndose a la tabla de signos positivos y negativos del diseño 2^5, como en Davies [36]. La inspección de esta tabla revela que las combinaciones de los tratamientos se asignan a los bloques de la siguiente manera:

Combinaciones de los tratamientos en el	Signo de *ADE*	Signo de *BCE*	Signo de *ABCD*
Bloque 1	−	−	+
Bloque 2	+	−	−
Bloque 3	−	+	−
Bloque 4	+	+	+

Observe que el producto de los signos de dos efectos cualesquiera de un bloque particular (por ejemplo *ADE* y *BCE*) produce el signo del otro efecto de ese bloque (en este caso, *ABCD*). Por lo tanto, *ADE*, *BCE* y *ABCD* están confundidos con los bloques.

Las propiedades de la teoría de grupos del bloque principal mencionadas en la sección 7-4 siguen siendo válidas. Por ejemplo, se observa que el producto de dos combinaciones de tratamientos del bloque principal produce otro elemento del bloque principal. Es decir,

$$ad \cdot bc = abcd \quad \text{y} \quad abe \cdot bde = ab^2de^2 = ad$$

etcétera. Para construir otro bloque se selecciona una combinación de tratamientos que no esté en el bloque principal (por ejemplo b), y b se multiplica por todas las combinaciones de tratamientos del bloque principal. Se obtiene así

$$b \cdot (1) = b \qquad b \cdot ad = abd \qquad b \cdot bc = b^2c = c \qquad b \cdot abcd = ab^2cd = acd$$

etcétera, lo que producirá las ocho combinaciones de tratamientos del bloque 3. En la práctica, el bloque principal puede obtenerse a partir de la definición de contrastes y de la propiedad de la teoría de grupos, y los demás bloques pueden determinarse a partir de estas combinaciones de los tratamientos aplicando el método que se presentó anteriormente.

El procedimiento general para construir un diseño 2^k confundido en cuatro bloques consiste en elegir dos efectos para generar los bloques, confundiéndose automáticamente un tercer efecto que es la interacción generalizada de las dos primeras. Después se construye el diseño utilizando las dos definiciones de contrastes (L_1, L_2) y las propiedades de la teoría de grupos del bloque principal. Al seleccionar los efectos que van a confundirse con los bloques, debe tenerse cuidado de obtener un diseño en el que no estén confundidos efectos que pueden ser de interés. Por ejemplo, en un diseño 2^5 podría elegirse confundir *ABCDE* y *ABD*, con lo cual se confunde automáticamente *CE*, un efecto que es de posible interés. Una mejor elección es confundir *ADE* y *BCE*, con lo cual se confunde automáticamente *ABCD*. Es preferible sacrificar información en las interacciones de tres factores *ADE* y *BCE* en lugar de la interacción de dos factores *CE*.

7-6 CONFUSIÓN DEL DISEÑO FACTORIAL 2^k EN 2^p BLOQUES

Los métodos descritos antes pueden extenderse a la construcción de un diseño factorial 2^k confundido (o mezclado) en 2^p bloques ($p < k$), donde cada bloque contiene exactamente 2^{k-p} corridas. Se seleccionan p efectos independientes que van a confundirse, donde por "independientes" se entiende que ninguno de los efectos elegidos es la interacción generalizada de los demás. Los bloques pueden generarse mediante

Tabla 7-8 Disposiciones de los bloques sugeridas para el diseño factorial 2^k

Número de factores, k	Número de bloques, 2^p	Tamaño del bloque, 2^{k-p}	Efectos elegidos para generar los bloques	Interacciones confundidas con los bloques
3	2	4	*ABC*	*ABC*
	4	2	*AB, AC*	*AB, AC, BC*
4	2	8	*ABCD*	*ABCD*
	4	4	*ABC, ACD*	*ABC, ACD, BD*
	8	2	*AB, BC, CD*	*AB, BC, CD, AC, BD, AD, ABCD*
5	2	16	*ABCDE*	*ABCDE*
	4	8	*ABC, CDE*	*ABC, CDE, ABDE*
	8	4	*ABE, BCE, CDE*	*ABE, BCE, CDE, AC, ABCD, BD, ADE*
	16	2	*AB, AC, CD, DE*	Todas las interacciones de dos y cuatro factores (15 efectos)
6	2	32	*ABCDEF*	*ABCDEF*
	4	16	*ABCF, CDEF*	*ABCF, CDEF, ABDE*
	8	8	*ABEF, ABCD, ACE*	*ABEF, ABCD, ACE, BCF, BDE, CDEF, ADF*
	16	4	*ABF, ACF, BDF, DEF*	*ABE, ACF, BDF, DEF, BC, ABCD, ABDE, AD, ACDE, CE, BDF, BCDEF, ABCEF, AEF, BE*
	32	2	*AB, BC, CD, DE, EF*	Todas las interacciones de dos, cuatro y seis factores (31 efectos)
7	2	64	*ABCDEFG*	*ABCDEFG*
	4	32	*ABCFG, CDEFG*	*ABCFG, CDEFG, ABDE*
	8	16	*ABC, DEF, AFG*	*ABC, DEF, AFG, ABCDEF, BCFG, ADEG, BCDEG*
	16	8	*ABCD, EFG, CDE, ADG*	*ABCD, EFG, CDE, ADG, ABCDEFG, ABE, BCG, CDFG, ADEF, ACEG, ABFG, BCEF, BDEG, ACF, BDF*
	32	4	*ABG, BCG, CDG, DEG, EFG*	*ABG, BCG, CDG, DEG, EFG, AC, BD, CE, DF, AE, BE, ABCD, ABDE, ABEF, BCDE, BCEF, CDEF, ABCDEFG, ADG, ACDEG, ACEFG, ABDFG, ABCEG, BEG, BDEFG, CFG, ADEF, ACDF, ABCF, AFG*
	64	2	*AB, BC, CD, DE, EF, FG*	Todas las interacciones de dos, cuatro y seis factores (63 efectos)

el uso de las p definiciones de contrastes $L_1, L_2, ..., L_p$ asociadas con estos efectos. Asimismo, se confundirán otros $2^p - p - 1$ efectos con los bloques, siendo éstos las interacciones generalizadas de los p efectos independientes elegidos inicialmente. Deberá tenerse cuidado al seleccionar los efectos que van a confundirse para que no se sacrifique información sobre los efectos que pueden ser de interés potencial.

El análisis estadístico de estos diseños es directo. Las sumas de cuadrados de todos los efectos se calculan como si no se hubiera hecho la formación de bloques. Después, la suma de cuadrados de los bloques se encuentra sumando las sumas de cuadrados de todos los efectos confundidos con los bloques.

Obviamente, la elección de los p efectos usados para generar el bloque es crítica, ya que la estructura de la confusión (o mezclado) del diseño depende directamente de ellos. En la tabla 7-8 se presenta una lista de diseños útiles. Para ilustrar el uso de esta tabla, suponga que quiere construirse un diseño 2^6 confundido en $2^3 = 8$ bloques con $2^3 = 8$ corridas cada uno. La tabla 7-8 indica que se elegirían $ABEF$, $ABCD$ y ACE como los $p = 3$ efectos independientes para generar los bloques. Los $2^p - p - 1 = 2^3 - 3 - 1 = 4$ efectos restantes que están confundidos son las interacciones generalizadas de estos tres; es decir,

$$(ABEF)(ABCD) = A^2B^2CDEF = CDEF$$
$$(ABEF)(ACE) = A^2BCE^2F = BCF$$
$$(ABCD)(ACE) = A^2BC^2ED = BDE$$
$$(ABEF)(ABCD)(ACE) = A^3B^2C^2DE^2F = ADF$$

En el problema 7-11 se le pide al lector que genere los ocho bloques de este diseño.

7-7 CONFUSIÓN PARCIAL

En la sección 7-4 se subrayó que, a menos que los experimentadores cuenten con una estimación previa del error o que estén dispuestos a suponer que ciertas interacciones son insignificantes, deben hacer réplicas del diseño para obtener una estimación del error. En la figura 7-3 se muestra un diseño factorial 2^3 en dos bloques con ABC confundido, con cuatro réplicas. Por el análisis de varianza de este diseño, el cual se presenta en la tabla 7-5, se observa que no puede sacarse información acerca de la interacción ABC debido a que ABC está confundido con los bloques *en todas las réplicas*. Se dice que este diseño está **completamente confundido** (o **mezclado**).

Considere la alternativa que se presenta en la figura 7-6. De nueva cuenta hay cuatro réplicas del diseño 2^3, pero en cada réplica se ha confundido una interacción *diferente*. Es decir, ABC está confundido en la réplica I, AB está confundido en la réplica II, BC está confundido en la réplica III y AC está confundido en la réplica IV. Como resultado puede obtenerse información de ABC a partir de los datos de las réplicas II, III y IV; información de AB puede obtenerse de las réplicas I, III y IV; información de AC puede obte-

Réplica I *ABC* Confundido		Réplica II *AB* Confundido		Réplica III *BC* Confundido		Réplica IV *AC* Confundido	
(1)	*a*	(1)	*a*	(1)	*b*	(1)	*a*
ab	*b*	*c*	*b*	*a*	*c*	*b*	*c*
ac	*c*	*ab*	*ac*	*bc*	*ab*	*ac*	*ab*
bc	*abc*	*abc*	*bc*	*abc*	*ac*	*abc*	*bc*

Figura 7-6 Confusión parcial en el diseño 2^3.

Tabla 7-9 Análisis de varianza de un diseño 2^3 parcialmente confundido

Fuente de variación	Grados de libertad
Réplicas	3
Bloques dentro de réplicas [o *ABC* (rép. I) + *AB* (rép. II) + *BC* (rép. III) + *AC* (rép. IV)]	4
A	1
B	1
C	1
AB (de las réplicas I, III y IV)	1
AC (de las réplicas I, II y III)	1
BC (de las réplicas I, II y IV)	1
ABC (de las réplicas I, III y IV)	1
Error	17
Total	31

nerse de las réplicas I, II y III; e información de *BC* puede obtenerse de las réplicas I, II y IV. Se dice que pueden obtenerse tres cuartas partes de la información de las interacciones porque no están confundidas en sólo tres réplicas. Yates [113b] llama a la relación 3/4 la **información relativa de los efectos confundidos.** Se dice que este diseño está **parcialmente confundido** (o **mezclado).**

En la tabla 7-9 se muestra el análisis de varianza de este diseño. Para calcular las sumas de cuadrados de las interacciones, sólo se usan los datos de las réplicas en las que no está confundida una interacción. La suma de cuadrados del error consta de las sumas de cuadrados de réplicas × sumas de cuadrados de efecto principal, más las sumas de cuadrados de réplicas × sumas de cuadrados de interacción para cada réplica en la que esa interacción no está confundida (por ejemplo, réplicas × *ABC* para las réplicas II, III y IV). Además, hay siete grados de libertad entre los ocho bloques. Es común hacer la partición de tres grados de libertad para las réplicas y cuatro grados de libertad para los bloques dentro de las réplicas. La composición de la suma de cuadrados de los bloques se muestra en la tabla 7-9 y se sigue directamente de la elección del efecto confundido en cada réplica.

EJEMPLO 7-3 ....

Un diseño 2^3 con confusión parcial

Considere el ejemplo 6-1, en el que se realizó un estudio para determinar el efecto del porcentaje de carbonatación (*A*), la presión de operación (*B*) y la velocidad de línea (*C*) sobre la altura de llenado de una bebida carbonatada. Suponga que cada lote de jarabe alcanza sólo para probar cuatro combinaciones de tratamientos. Por lo tanto, cada réplica del diseño 2^3 debe correrse en dos bloques. Se corren dos réplicas, con *ABC* confundido en la réplica I y *AB* confundido en la réplica II. Los datos son los siguientes:

Réplica I, *ABC* confundido: Bloque 1	Réplica I, *ABC* confundido: Bloque 2	Réplica II, *AB* confundido: Bloque 1	Réplica II, *AB* confundido: Bloque 2
(1) = −3	*a* = 0	(1) = −1	*a* = 1
ab = 2	*b* = −1	*c* = 0	*b* = 0
ac = 2	*c* = −1	*ab* = 3	*ac* = 1
bc = 1	*abc* = 6	*abc* = 5	*bc* = 1

Tabla 7-10 Análisis de varianza del ejemplo 7-3

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0	Valor P
Réplicas	1.00	1	1.00	-	
Bloques dentro de las réplicas	2.50	2	1.25	–	
A	36.00	1	36.00	48.00	0.0001
B	20.25	1	20.25	27.00	0.0035
C	12.25	1	12.25	16.33	0.0099
AB (sólo en la réplica I)	0.50	1	0.50	0.67	0.4503
AC	0.25	1	0.25	0.33	0.5905
BC	1.00	1	1.00	1.33	0.3009
ABC (sólo en la réplica II)	0.50	1	0.50	0.67	0.4503
Error	3.75	5	0.75	–	
Total	78.00	15	–	–	

Las sumas de cuadrados de A, B, C, AC y BC pueden calcularse de la manera usual, utilizando las 16 observaciones. Sin embargo, SS_{ABC} debe encontrarse utilizando únicamente los datos de la réplica II y SS_{AB} utilizando únicamente los datos de la réplica I de la siguiente manera:

$$SS_{ABC} = \frac{[a+b+c+abc-ab-ac-bc-(1)]^2}{n2^k}$$

$$= \frac{[1+0+0+5-3-1-1-(-1)]^2}{(1)(8)} = 0.50$$

$$S_{AB} = \frac{[(1)+abc-ac+c-a-b+ab-bc]^2}{n2^k}$$

$$= \frac{[-3+6-2+(-1)-0-(-1)+2-1]^2}{(1)(8)} = 0.50$$

La suma de cuadrados de las réplicas es, en general,

$$SS_{\text{Rep}} = \sum_{h=1}^{n} \frac{R_h^2}{2^k} - \frac{y_{...}^2}{N}$$

$$= \frac{(6)^2 + (10)^2}{8} - \frac{(16)^2}{16} = 1.00$$

donde R_h es el total de las observaciones en la réplica h-ésima. La suma de cuadrados de los bloques es la suma de SS_{ABC} de la réplica I y SS_{AB} de la réplica II, o $SS_{\text{Bloques}} = 2.50$.

En la tabla 7-10 se resume el análisis de varianza. Los tres efectos principales son importantes.

7-8 PROBLEMAS

7-1. Considere el experimento descrito en el problema 6-1. Analizar este experimento suponiendo que cada réplica representa un bloque de un solo turno de producción.

7-2. Considere el experimento descrito en el problema 6-5. Analizar este experimento suponiendo que cada una de las cuatro réplicas representa un bloque.

7-3. Considere el experimento de la formación de fisuras en la aleación de níquel y titanio descrito en el problema 6-15. Suponga que sólo pudieron hacerse 16 corridas en un solo día, por lo que cada réplica se trató como un bloque. Analizar el experimento y sacar conclusiones.

7-4. Considere los datos de la primera réplica del problema 6-1. Suponga que no fue posible correr todas estas observaciones utilizando barras del mismo lote. Establecer un diseño para correr estas observaciones en dos bloques de cuatro observaciones cada uno con ABC confundido. Analizar los datos.

7-5. Considere los datos de la primera réplica del problema 6-7. Construir un diseño con dos bloques de ocho observaciones cada uno con $ABCD$ confundido. Analizar los datos.

7-6. Repetir el problema 7-5 suponiendo que se requieren cuatro bloques. Confundir ABD y ABC (y por consiguiente CD) con los bloques.

7-7. Utilizando los datos del diseño 2^5 del problema 6-21, construir y analizar un diseño en dos bloques con $ABCDE$ confundido con los bloques.

7-8. Repetir el problema 7-7 suponiendo que se necesitan cuatro bloques. Sugerir un esquema de confusión (o mezclado) razonable.

7-9. Considere los datos del diseño 2^5 del problema 6-21. Suponga que fue necesario correr este diseño en cuatro bloques con $ACDE$ y BCD (y por consiguiente ABE) confundidos. Analizar los datos de este diseño.

7-10. Diseñar un experimento para confundir un diseño factorial 2^6 en cuatro bloques. Sugerir un esquema de confusión apropiado, diferente del que se ilustró en la tabla 7-8.

7-11. Considere el diseño 2^6 en ocho bloques con ocho corridas cada uno con $ABCD$, ACE y $ABEF$ como los efectos independientes elegidos para confundirlos con los bloques. Generar el diseño. Encontrar los demás efectos confundidos con los bloques.

7-12. Considere el diseño 2^2 en dos bloques con AB confundido. Hacer la demostración algebraica de que $SS_{AB} = SS_{\text{Bloques}}$.

7-13. Considere los datos del ejemplo 7-2. Suponga que todas las observaciones del bloque 2 se incrementan en 20. Analizar los datos que resultarían. Estimar el efecto de bloque. ¿Puede el lector explicar su magnitud? ¿Los bloques parecen ser ahora un factor importante? ¿Hay otras estimaciones de los efectos que sufran el impacto de este cambio hecho en los datos?

7-14. Suponga que en el problema 6-1 se confundió ABC en la réplica I, AB en la réplica II y BC en la réplica III. Calcular las estimaciones de los efectos. Construir la tabla del análisis de varianza.

7-15. Repetir el problema 6-1 suponiendo que ABC se confundió con los bloques en todas las réplicas.

7-16. Suponga que en el problema 6-7 $ABCD$ se confundió en la réplica I y ABC se confundió en la réplica II. Realizar el análisis estadístico de este diseño.

7-17. Construir un diseño 2^3 con ABC confundido en las dos primeras réplicas y BC confundido en la tercera réplica. Delinear el análisis de varianza y comentar la información obtenida.

8 Diseños factoriales fraccionados de dos niveles

8-1 INTRODUCCIÓN

Cuando el número de factores de un diseño factorial 2^k se incrementa, el número de corridas necesarias para realizar una réplica completa del diseño rebasa con rapidez los recursos de la mayoría de los experimentadores. Por ejemplo, una réplica completa de un diseño 2^6 requiere 64 corridas. En este diseño, sólo 6 de los 63 grados de libertad corresponden a los efectos principales, y sólo 15 a las interacciones de dos factores. Los 42 grados de libertad restantes se asocian con las interacciones de tres o más factores.

Si el experimentador puede suponer razonablemente que ciertas interacciones de orden superior son insignificantes, es posible obtener información de los efectos principales y las interacciones de orden inferior corriendo únicamente una fracción del experimento factorial completo. Estos **diseños factoriales fraccionados** se encuentran entre los tipos de diseños de uso más generalizado en el diseño de productos y procesos y en el mejoramiento de procesos.

Una de las principales aplicaciones de los diseños factoriales fraccionados es en los **experimentos de tamizado** o **exploración.** Se trata de experimentos en los que se consideran muchos factores y el objetivo es identificar aquellos factores (en caso de haberlos) que tienen efectos grandes. Los experimentos de tamizado suelen realizarse en las etapas iniciales de un proyecto, cuando es posible que muchos de los factores considerados en un principio tengan un efecto reducido o nulo sobre la respuesta. Entonces los factores que se identifican como importantes se investigan con mayor detalle en experimentos subsecuentes.

El uso exitoso de los diseños factoriales fraccionados se basa en tres ideas clave:

1. ***El principio de efectos esparcidos*** o ***escasez de efectos.*** Cuando hay varias variables, es posible que el sistema o proceso esté dominado principalmente por algunos de los efectos principales y las interacciones de orden inferior.
2. ***La propiedad de proyección.*** Los diseños factoriales fraccionados pueden proyectarse en diseños más fuertes (más grandes) en el subconjunto de los factores significativos.
3. ***Experimentación secuencial.*** Es posible combinar las corridas de dos (o más) diseños factoriales fraccionados para ensamblar secuencialmente un diseño más grande para estimar los efectos de los factores y las interacciones de interés.

Este capítulo se enfoca en estos principios, los cuales se ilustran con varios ejemplos.

8-2 LA FRACCIÓN UN MEDIO DEL DISEÑO 2^k

Considere una situación en la que tres factores, cada uno con dos niveles, son de interés, pero los experimentadores no están en posición de correr las $2^3 = 8$ combinaciones de tratamientos. Sin embargo, pueden llevar a cabo cuatro corridas. Esto sugiere una fracción un medio de un diseño 2^3. Puesto que el diseño contiene $2^{3-1} = 4$ combinaciones de tratamientos, es común llamar **diseño 2^{3-1}** a una fracción un medio del diseño 2^3.

En la tabla 8-1 se muestra la agrupación de signos positivos y negativos del diseño 2^3. Suponga que se seleccionan las cuatro combinaciones de tratamientos *a, b, c* y *abc* como la fracción un medio con la que se trabajará. Estas corridas se muestran en la parte superior de la tabla 8-1 y en la figura 8-1*a*.

Observe que el diseño 2^{3-1} se forma seleccionando sólo las combinaciones de tratamientos que tienen signo positivo en la columna *ABC*. Por lo tanto, a *ABC* se le llama el **generador** de esta fracción particular. En ocasiones se hará referencia a un generador, por ejemplo *ABC*, como una **palabra.** Además, la columna identidad *I* también es siempre positiva, por lo que a

$$I = ABC$$

se le llama la **relación de definición** del diseño. En general, la relación de definición de un diseño factorial fraccionado será siempre el conjunto de todas las columnas que son iguales a la columna identidad *I*.

Las combinaciones de tratamientos del diseño 2^{3-1} producen tres grados de libertad que pueden usarse para estimar los efectos principales. Con referencia a la tabla 8-1, se observa que las combinaciones lineales de las observaciones usadas para estimar los efectos principales de *A, B* y *C* son

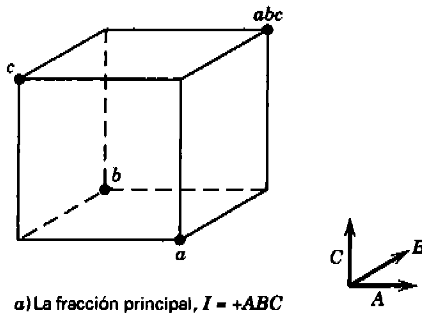
$$\ell_A = \tfrac{1}{2}(a - b - c + abc)$$
$$\ell_B = \tfrac{1}{2}(-a + b - c + abc)$$
$$\ell_C = \tfrac{1}{2}(-a - b + c + abc)$$

También es sencillo verificar que las combinaciones lineales de las observaciones usadas para estimar las interacciones de dos factores son

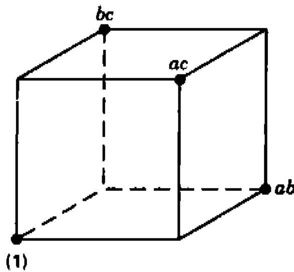
$$\ell_{BC} = \tfrac{1}{2}(a - b - c + abc)$$
$$\ell_{AC} = \tfrac{1}{2}(-a + b - c + abc)$$
$$\ell_{AB} = \tfrac{1}{2}(-a - b + c + abc)$$

Tabla 8-1 Signos positivos y negativos del diseño factorial 2^3

Combinación de tratamientos	Efecto factorial							
	I	*A*	*B*	*C*	*AB*	*AC*	*BC*	*ABC*
a	+	+	−	−	−	−	+	+
b	+	−	+	−	−	+	−	+
c	+	−	−	+	+	−	−	+
abc	+	+	+	+	+	+	+	+
ab	+	+	+	−	+	−	−	−
ac	+	+	−	+	−	+	−	−
bc	+	−	+	+	−	−	+	−
(1)	+	−	−	−	+	+	+	−



a) La fracción principal, $I = +ABC$



b) La fracción alterna, $I = -ABC$

Figura 8-1 Las dos fracciones un medio del diseño 2^3.

Por lo tanto, $\ell_A = \ell_{BC}$, $\ell_B = \ell_{AC}$ y $\ell_C = \ell_{AB}$; por consiguiente, es imposible diferenciar entre A y BC, entre B y AC y entre C y AB. De hecho, cuando se estiman A, B y C, se están estimando *en realidad* $A + BC$, $B + AC$ y $C + AB$. A dos o más efectos que tienen esta propiedad se les llama **alias.** En el ejemplo tratado aquí, A y BC son alias, B y AC son alias y C y AB son alias. Esto se indica con la notación $\ell_A \rightarrow A + BC$, $\ell_B \rightarrow B + AC$ y $\ell_C \rightarrow C + AB$.

La estructura de los alias para este diseño puede determinarse con facilidad utilizando la relación de definición $I = ABC$. Al multiplicar cualquier columna (o efecto) por la relación de definición se obtienen los alias de esa columna (o efecto). En el ejemplo tratado aquí se encuentra que el alias de A es

$$A \cdot I = A \cdot ABC = A^2 BC$$

o, puesto que el cuadrado de cualquier columna es la identidad I,

$$A = BC$$

De manera similar, se encuentra que los alias de B y C son

$$B \cdot I = B \cdot ABC$$
$$B = AB^2C = AC$$

y

$$C \cdot I = C \cdot ABC$$
$$C = ABC^2 = AB$$

A esta fracción un medio, con $I = +ABC$, suele llamársele la **fracción principal.**

Suponga ahora que se eligió la *otra* fracción un medio, es decir, las combinaciones de tratamientos de la tabla 8-1 asociadas con los signos negativos de la columna ABC. Esta fracción un medio **alterna o com-**

plementaria (la cual se compone de las corridas (1), *ab, ac* y *bc*) se ilustra en la figura 8-1*b*. La relación de definición de este diseño es

$$I = -ABC$$

De la combinación lineal de las observaciones, por ejemplo ℓ'_A, ℓ'_B y ℓ'_C, de la fracción alterna se obtiene

$$\ell'_A \rightarrow A - BC$$
$$\ell'_B \rightarrow B - AC$$
$$\ell'_C \rightarrow C - AB$$

Por lo tanto, cuando se estiman *A, B* y *C* con esta fracción particular, en realidad se están estimando *A–BC, B–AC* y *C–AB*.

En la práctica, no importa cuál de las fracciones se usa. Ambas fracciones pertenecen a la misma **familia;** es decir, las dos fracciones un medio forman un diseño 2^3 completo. Esto puede observarse con facilidad con referencia a los incisos *a* y *b* de la figura 8-1.

Suponga que después de correr una de las fracciones un medio del diseño 2^3, también se corrió la otra. Por lo tanto, se cuenta ahora con las ocho corridas asociadas con el diseño 2^3 completo. Pueden obtenerse entonces las estimaciones sin alias de todos los efectos analizando las ocho corridas como un diseño 2^3 completo en dos bloques de cuatro corridas cada uno. Esto también podría hacerse sumando y restando la combinación lineal de los efectos de las dos fracciones individuales. Por ejemplo, considere $\ell_A \rightarrow A + BC$ y $\ell'_A \rightarrow A - BC$. Esto implica que

$$\tfrac{1}{2}(\ell_A + \ell'_A) = \tfrac{1}{2}(A + BC + A - BC) \rightarrow A$$

y que

$$\tfrac{1}{2}(\ell_A - \ell'_A) = \tfrac{1}{2}(A + BC - A + BC) \rightarrow BC$$

Por lo tanto, para los tres pares de combinaciones lineales se obtendría lo siguiente:

i	De $\frac{1}{2}(\ell_i + \ell'_i)$	De $\frac{1}{2}(\ell_i - \ell'_i)$
A	*A*	*BC*
B	*B*	*AC*
C	*C*	*AB*

Resolución del diseño

Al diseño 2^{3-1} precedente se le llama **diseño de resolución III.** En este diseño, los efectos principales son alias de las interacciones de dos factores. Un diseño es de resolución *R* cuando ningún efecto del factor *p* es alias de otro efecto que contiene menos de $R-p$ factores. Por lo general se emplea un subíndice con un numeral romano para denotar la resolución del diseño; por lo tanto, la fracción un medio del diseño 2^3 con la relación de definición $I = ABC$ (o $I = -ABC$) es un diseño 2^{3-1}_{III}.

Los diseños de resolución III, IV y V son particularmente importantes. A continuación se presentan las definiciones de estos diseños y un ejemplo de cada uno:

1. ***Diseños de resolución III.*** Se trata de diseños en los que ninguno de los efectos principales es alias de ningún otro efecto principal, pero los efectos principales son alias de las interacciones de dos factores, y algunas de las interacciones de dos factores pueden ser alias entre sí. El diseño 2^{3-1} de la tabla 8-1 es un diseño de resolución III (2^{3-1}_{III}).
2. ***Diseños de resolución IV.*** Se trata de diseños en los que ninguno de los efectos principales es alias de ningún otro efecto principal *ni* de las interacciones de dos factores, pero las interacciones de

dos factores son alias entre sí. Un diseño 2^{4-1} con $I = ABCD$ es un diseño de resolución IV (2_{IV}^{4-1}).

3. ***Diseños de resolución V.*** Se trata de diseños en los que ninguno de los efectos principales ni de las interacciones de dos factores son alias de otro efecto principal o interacción de dos factores, pero las interacciones de dos factores son alias de las interacciones de tres factores. Un diseño 2^{5-1} con $I = ABCDE$ es un diseño de resolución V (2_V^{5-1}).

En general, la resolución de un diseño factorial fraccionado de dos niveles es igual al menor número de letras en cualquier palabra de la relación de definición. Por consiguiente, los diseños precedentes podrían denominarse diseños de tres, cuatro y cinco letras, respectivamente. Por lo común, es preferible emplear diseños fraccionados que tengan la resolución más alta posible que sea consistente con el grado de fraccionamiento requerido. Entre más alta sea la resolución, menos restrictivos serán los supuestos que se requieren respecto de cuáles de las interacciones son insignificantes para obtener una interpretación única de los datos.

Construcción de fracciones un medio

Una fracción un medio del diseño 2^k de la resolución más alta puede construirse apuntando el **diseño básico,** que consta de las corridas de un diseño factorial 2^{k-1} *completo*, y agregándole después el factor k-ésimo identificando sus niveles positivo y negativo con los signos positivo y negativo de la interacción $ABC \cdots (K-1)$ del orden más alto. Por lo tanto, el diseño factorial fraccionado 2_{III}^{3-1} se obtiene apuntando el diseño 2^2 completo como diseño básico e igualando después el factor C con la interacción AB. La fracción alterna se obtendría igualando el factor C con la interacción $-AB$. Este enfoque se ilustra en la tabla 8-2. Observe que el diseño básico siempre tiene el número correcto de corridas (renglones), pero le falta una columna. El generador $I = ABC \cdots K$ se resuelve entonces para la columna faltante (K), de tal modo que $K = ABC \cdots (K-1)$ define el producto de los signos positivos y negativos que deberá usarse en cada renglón para producir los niveles del factor k-ésimo.

Observe que podría usarse *cualquier* efecto de interacción para generar la columna del factor k-ésimo. Sin embargo, al utilizarse cualquier efecto que no sea $ABC \cdots (K-1)$, no se producirá el diseño con la resolución más alta posible.

Otra forma de visualizar la construcción de una fracción un medio es mediante la partición de las corridas en dos bloques con la interacción de orden más alto $ABC \cdots K$ confundida. Cada bloque es un diseño factorial fraccionado 2^{k-1} con la resolución más alta.

Proyección de fracciones en diseños factoriales

Cualquier diseño factorial fraccionado de resolución R contiene diseños factoriales completos (posiblemente diseños factoriales con réplicas) en cualquier subconjunto de $R-1$ factores. Éste es un concepto importante y útil. Por ejemplo, si un experimentador tiene varios factores de interés potencial pero piensa

Tabla 8-2 Las dos fracciones un medio del diseño 2^3

	Diseño factorial 2^2 completo (diseño básico)		2_{III}^{3-1}, $I = ABC$			2_{III}^{3-1}, $I = -ABC$		
Corrida	A	B	A	B	$C = AB$	A	B	$C = -AB$
1	−	−	−	−	+	−	−	−
2	+	−	+	−	−	+	−	+
3	−	+	−	+	−	−	+	+
4	+	+	+	+	+	+	+	−

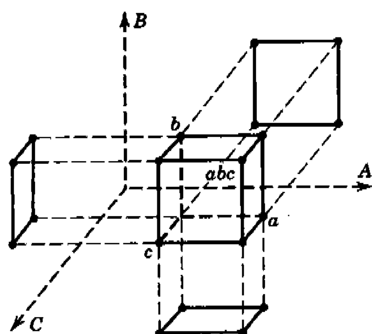


Figura 8-2 Proyección de un diseño 2_{III}^{3-1} en tres diseños 2^2.

que sólo $R-1$ de ellos tienen efectos importantes, entonces un diseño factorial fraccionado de resolución R es la elección de diseño apropiada. Si el experimentador está en lo correcto, el diseño factorial fraccionado de resolución R se proyectará en un diseño factorial completo en los $R-1$ factores significativos. Este proceso se ilustra en la figura 8-2 para el diseño 2_{III}^{3-1}, el cual se proyecta en un diseño 2^2 en cada subconjunto de dos factores.

Puesto que la máxima resolución posible de una fracción un medio del diseño 2^k es $R = k$, todos los diseños 2^{k-1} se proyectarán en un factorial completo en $(k-1)$ cualquiera de los k factores originales. Además, un diseño 2^{k-1} puede proyectarse en dos réplicas de un factorial completo en cualquier subconjunto de $k-2$ factores, cuatro réplicas de un factorial completo en cualquier subconjunto de $k-3$ factores, etcétera.

EJEMPLO 8-1 ··························

Considere el experimento del índice de filtración del ejemplo 6-2. El diseño original, ilustrado en la tabla 6-10, es una sola réplica del diseño 2^4. En ese ejemplo se encontró que los efectos principales de A, C y D y las interacciones AC y AD eran diferentes de cero. Se retoma ahora este experimento y se simula lo que habría ocurrido si se hubiera corrido una fracción un medio del diseño 2^4 en vez del diseño factorial completo.

Se usará el diseño 2^{4-1} con $I = ABCD$, ya que esta elección del generador dará como resultado un diseño con la resolución más alta posible (IV). Para construir el diseño, primero se apunta el diseño básico, el cual es un diseño 2^3, como se muestra en las tres primeras columnas de la tabla 8-3. Este diseño básico tiene el número necesario de corridas (ocho) pero sólo tres columnas (factores). Para encontrar los niveles del cuarto factor, se resuelve $I = ABCD$ para D, o $D = ABC$. Por lo tanto, el nivel de D de cada corrida

Tabla 8-3 El diseño 2_{IV}^{4-1} con la relación de definición $I=ABCD$

	Diseño básico					
Corrida	A	B	C	$D = ABC$	Combinación de tratamientos	Índice de filtración
1	−	−	−	−	(1)	45
2	+	−	−	+	*ad*	100
3	−	+	−	+	*bd*	45
4	+	+	−	−	*ab*	65
5	−	−	+	+	*cd*	75
6	+	−	+	−	*ac*	60
7	−	+	+	−	*bc*	80
8	+	+	+	+	*abcd*	96

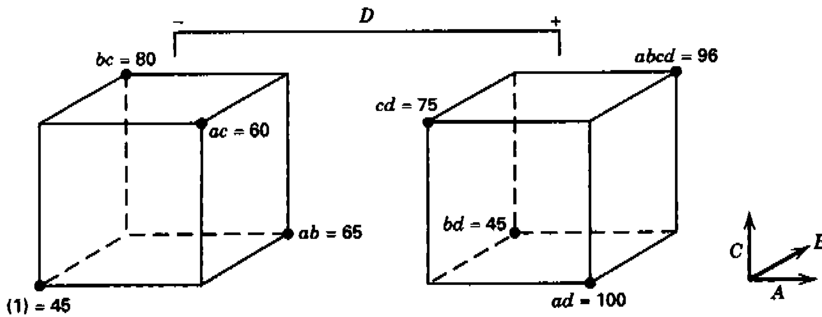


Figura 8-3 El diseño 2_{IV}^{4-1} para el experimento del índice de filtración del ejemplo 8-1.

es el producto de los signos positivos y negativos de las columnas A, B y C. El proceso se ilustra en la tabla 8-3. Puesto que el generador $ABCD$ es positivo, este diseño 2_{IV}^{4-1} es la fracción principal. El diseño se ilustra gráficamente en la figura 8-3.

Utilizando la relación de definición, se observa que cada uno de los efectos principales es alias de una interacción de tres factores; es decir, $A = A^2BCD = BCD$, $B = AB^2CD = ACD$, $C = ABC^2D = ABD$ y $D = ABCD^2 = ABC$. Además, cada interacción de dos factores es alias de otra interacción de dos factores. Estas relaciones de los alias son $AB = CD$, $AC = BD$ y $BC = AD$. Los cuatro efectos principales más los tres pares de alias de interacciones de dos factores representan los siete grados de libertad del diseño.

En este punto, normalmente se aleatorizarían las ocho corridas y se llevaría a cabo el experimento. Puesto que se ha corrido ya el diseño 2^4 completo, simplemente se seleccionan los ocho índices de filtración observados del ejemplo 6-2 que corresponden a las corridas del diseño 2_{IV}^{4-1}. Estas observaciones se muestran en la última columna de la tabla 8-3, así como en la figura 8-3.

En la tabla 8-4 se muestran las estimaciones de los efectos obtenidas de este diseño 2_{IV}^{4-1}. Para ilustrar los cálculos, la combinación lineal de las observaciones asociadas con el efecto de A es

$$\ell_A = \tfrac{1}{4}(-45+100-45+65-75+60-80+96) = 19.00 \rightarrow A + BCD$$

mientras que para el efecto AB se obtendría

$$\ell_{AB} = \tfrac{1}{4}(45-100-45+65+75-60-80+96) = -1.00 \rightarrow AB + CD$$

Por la inspección de la información de la tabla 8-4, no es irrazonable concluir que los efectos principales de A, C y D son grandes. Además, si A, C y D son los efectos principales importantes, entonces es lógico concluir que las dos cadenas de alias de interacciones $AC + BD$ y $AD + BC$ tienen efectos grandes, ya que

Tabla 8-4 Estimaciones de los efectos y los alias del ejemplo 8-1[a]

Estimación	Estructura de los alias
$\ell_A = 19.00$	$\ell_A \rightarrow A + BCD$
$\ell_B = 1.50$	$\ell_B \rightarrow B + ACD$
$\ell_C = 14.00$	$\ell_C \rightarrow C + ABD$
$\ell_D = 16.50$	$\ell_D \rightarrow D + ABC$
$\ell_{AB} = -1.00$	$\ell_{AB} \rightarrow AB + CD$
$\ell_{AC} = -18.50$	$\ell_{AC} \rightarrow AC + BD$
$\ell_{AD} = 19.00$	$\ell_{AD} \rightarrow AD + BC$

[a]Los efectos significativos se indican en negritas.

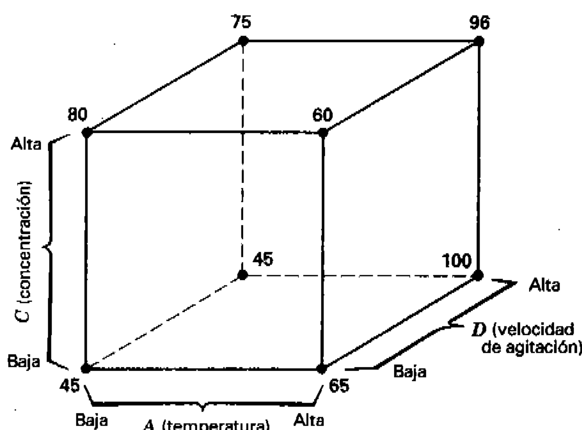


Figura 8-4 Proyección del diseño 2_{IV}^{4-1} en un diseño 2^3 en A, C y D para el ejemplo 8-1.

las interacciones AC y AD también son significativas. En otras palabras, si A, C y D son significativos, entonces lo más posible es que las interacciones significativas sean AC y AD. Se trata de una aplicación de la **navaja de Ockham** (en honor de Guillermo de Ockham), un principio científico que establece que cuando uno se confronta con varias interpretaciones posibles de un fenómeno, la interpretación más simple suele ser la correcta. Observe que esta interpretación concuerda con las conclusiones del análisis del diseño 2^4 completo del ejemplo 6-2.

Puesto que el factor B no es significativo, puede sacarse de consideración. Por consiguiente, este diseño 2_{IV}^{4-1} puede proyectarse en una sola réplica del diseño 2^3 en los factores A, C y D, como se muestra en la figura 8-4. El examen visual de esta gráfica de cubo nos hace sentirnos más cómodos con las conclusiones a las que se llegó antes. Observe que si la temperatura (A) está en el nivel bajo, la concentración (C) tiene un efecto positivo grande, mientras que si la temperatura está en el nivel alto, la concentración tiene un efecto muy pequeño. Esto se debe probablemente a una interacción AC. Además, si la temperatura está en el nivel bajo, el efecto de la velocidad de agitación (D) es insignificante, mientras que si la temperatura está en el nivel alto, la velocidad de agitación tiene un efecto positivo grande. Esto se debe probablemente a la interacción AD que se identificó de manera tentativa unos párrafos antes.

Con base en el análisis anterior, puede obtenerse ahora un modelo para predecir el índice de filtración en la región experimental. Este modelo es

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_3 x_3 + \hat{\beta}_4 x_4 + \hat{\beta}_{13} x_1 x_3 + \hat{\beta}_{14} x_1 x_4$$

donde x_1, x_3 y x_4 son variables codificadas ($-1 \le x_i \le +1$) que representan a A, C y D, y las $\hat{\beta}$ son coeficientes de regresión que pueden obtenerse a partir de las estimaciones de los efectos como se hizo anteriormente. Por lo tanto, la ecuación de predicción es

$$\hat{y} = 70.75 + \left(\frac{19.00}{2}\right) x_1 + \left(\frac{14.00}{2}\right) x_3 + \left(\frac{16.50}{2}\right) x_4 + \left(\frac{-18.50}{2}\right) x_1 x_3 + \left(\frac{19.00}{2}\right) x_1 x_4$$

Recuerde que la ordenada al origen $\hat{\beta}_0$ es el promedio de todas las respuestas en las ocho corridas del diseño. Este modelo es muy similar al que resultó del diseño factorial 2^k completo del ejemplo 6-2.

....

EJEMPLO 8-2 ....

Un diseño 2^{5-1} usado para mejorar un proceso

Se investigaron cinco factores en un proceso de manufactura de un circuito integrado en un diseño 2^{5-1} con el objetivo de mejorar el rendimiento del proceso. Los cinco factores fueron A = ajuste de apertura (pequeña, grande), B = tiempo de exposición (20% abajo del nominal, 20% arriba del nominal), C = tiempo de desarrollo (30 s, 45 s), D = tamaño de la máscara (pequeña, grande) y E = tiempo de grabado (14.5 min, 15.5 min). En la tabla 8-5 se muestra la construcción del diseño 2^{5-1}. Observe que el diseño se construyó apuntando el diseño básico que tiene 16 corridas (un diseño 2^4 en A, B, C y D), seleccionando $ABCDE$ como generador, y ajustando después los niveles del quinto factor $E = ABCD$. En la figura 8-5 se presenta una representación geométrica del diseño.

La relación de definición del diseño es $I = ABCDE$. Por consiguiente, todos los efectos principales son alias de una interacción de cuatro factores (por ejemplo, $\ell_A \rightarrow A + BCDE$), y cada una de las interacciones de dos factores son alias de una interacción de tres factores (por ejemplo, $\ell_{AB} \rightarrow AB + CDE$). Por lo tanto, el diseño es de resolución V. Se esperaría que este diseño 2^{5-1} proporcionara excelente información respecto de los efectos principales y las interacciones de dos factores.

La tabla 8-6 contiene las estimaciones de los efectos, las sumas de cuadrados y los coeficientes del modelo de regresión para los 15 efectos de este experimento. En la figura 8-6 se presenta la gráfica de probabilidad normal de las estimaciones de los efectos de este experimento. Los efectos principales de A, B y C y la interacción AB son grandes. Recuerde que, debido a los alias, estos efectos son en realidad $A + BCDE$, $B + ACDE$, $C + ABDE$ y $AB + CDE$. Sin embargo, puesto que parece plausible que las interacciones de tres factores y de órdenes superiores sean insignificantes, uno siente seguridad en concluir que sólo A, B, C y AB son los efectos importantes.

En la tabla 8-7 se resume el análisis de varianza de este experimento. La suma de cuadrados del modelo es $SS_{\text{Modelo}} = SS_A + SS_B + SS_C + SS_{AB} = 5747.25$, y esto explica más de 99% de la variabilidad total del rendimiento. En la figura 8-7 se presenta la gráfica de probabilidad normal de los residuales

Tabla 8-5 Un diseño 2^{5-1} para el ejemplo 8-2

Corrida	Diseño básico A	B	C	D	$E = ABCD$	Combinación de tratamientos	Rendimiento
1	−	−	−	−	+	*e*	8
2	+	−	−	−	−	*a*	9
3	−	+	−	−	−	*b*	34
4	+	+	−	−	+	*abe*	52
5	−	−	+	−	−	*c*	16
6	+	−	+	−	+	*ace*	22
7	−	+	+	−	+	*bce*	45
8	+	+	+	−	−	*abc*	60
9	−	−	−	+	−	*d*	6
10	+	−	−	+	+	*ade*	10
11	−	+	−	+	+	*bde*	30
12	+	+	−	+	−	*abd*	50
13	−	−	+	+	+	*cde*	15
14	+	−	+	+	−	*acd*	21
15	−	+	+	+	−	*bcd*	44
16	+	+	+	+	+	*abcde*	63

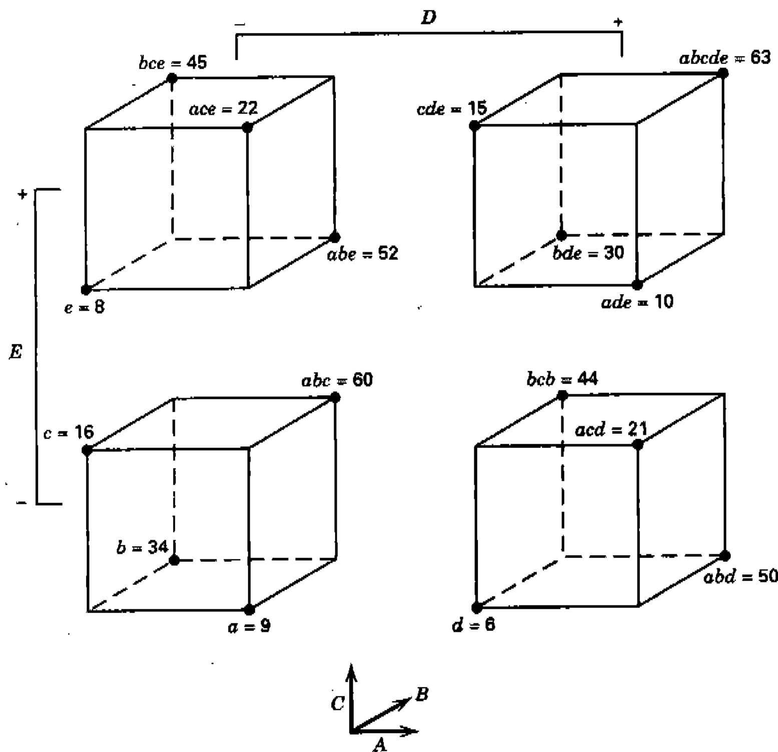


Figura 8-5 El diseño 2_V^{5-1} del ejemplo 8-2.

Tabla 8-6 Efectos, coeficientes de regresión y sumas de cuadrados del ejemplo 8-2

Variable	Nombre	Nivel −1	Nivel +1
A	Apertura	−1.000	1.000
B	Tiempo de desarrollo	−1.000	1.000
C	Tiempo de exposición	−1.000	1.000
D	Tamaño de la máscara	−1.000	1.000
E	Tiempo de grabado	−1.000	1.000

Variable	Coeficiente de regresión	Efecto estimado	Suma de cuadrados
Promedio global	30.3125		
A	5.5625	11.1250	495.062
B	16.9375	33.8750	4590.062
C	5.4375	10.8750	473.062
D	−0.4375	−0.8750	3.063
E	0.3125	0.6250	1.563
AB	3.4375	6.8750	189.063
AC	0.1875	0.3750	0.563
AD	0.5625	1.1250	5.063
AE	0.5625	1.1250	5.063
BC	0.3125	0.6250	1.563
BD	−0.0625	−0.1250	0.063
BE	−0.0625	−0.1250	0.063
CD	0.4375	0.8750	3.063
CE	0.1875	0.3750	0.563
DE	−0.6875	−1.3750	7.563

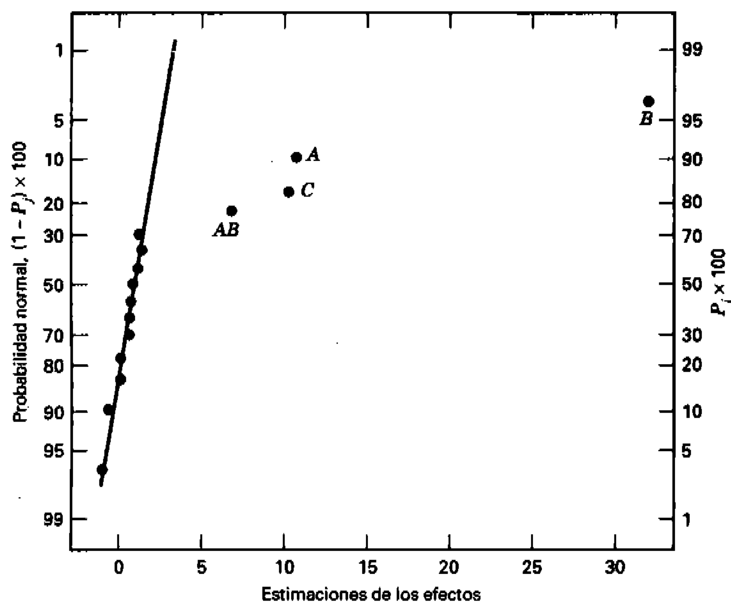


Figura 8-6 Gráfica de probabilidad normal de los efectos del ejemplo 8-2.

Tabla 8-7 Análisis de varianza del ejemplo 8-2

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0	Valor P
A (Apertura)	495.0625	1	495.0625	193.20	<0.0001
B (Tiempo de exposición)	4590.0625	1	4590.0625	1791.24	<0.0001
C (Tiempo de desarrollo)	473.0625	1	473.0625	184.61	<0.0001
AB	189.0625	1	189.0625	73.78	<0.0001
Error	28.1875	11	2.5625		
Total	5775.4375	15			

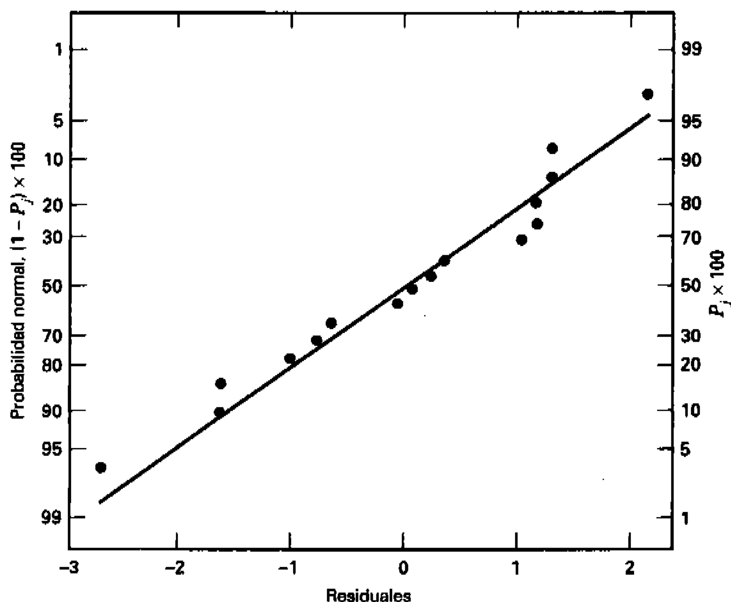


Figura 8-7 Gráfica de probabilidad normal de los residuales del ejemplo 8-2.

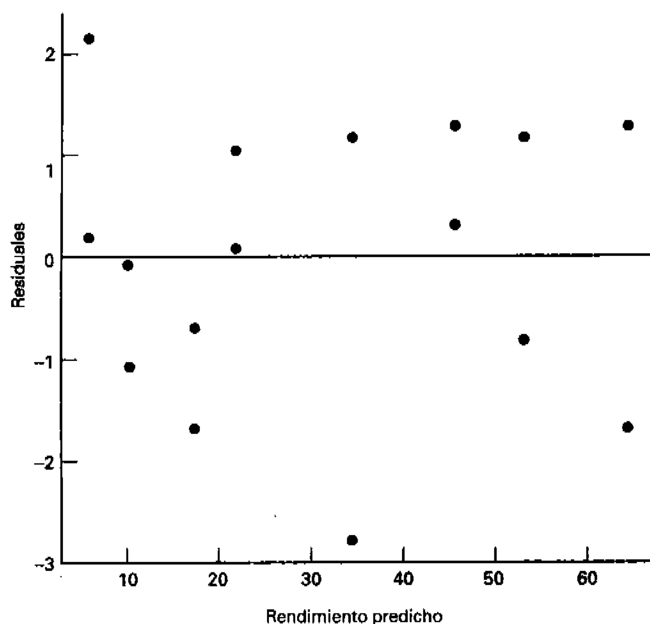


Figura 8-8 Gráfica de los residuales contra el rendimiento predicho para el ejemplo 8-2.

y la figura 8-8 es una gráfica de los residuales contra los valores predichos. Ambas gráficas son satisfactorias.

Los tres factores A, B y C tienen efectos positivos grandes. La interacción apertura-tiempo de exposición o AB se grafica en la figura 8-9. Esta gráfica confirma que el rendimiento es más alto cuando tanto A como B están en el nivel alto.

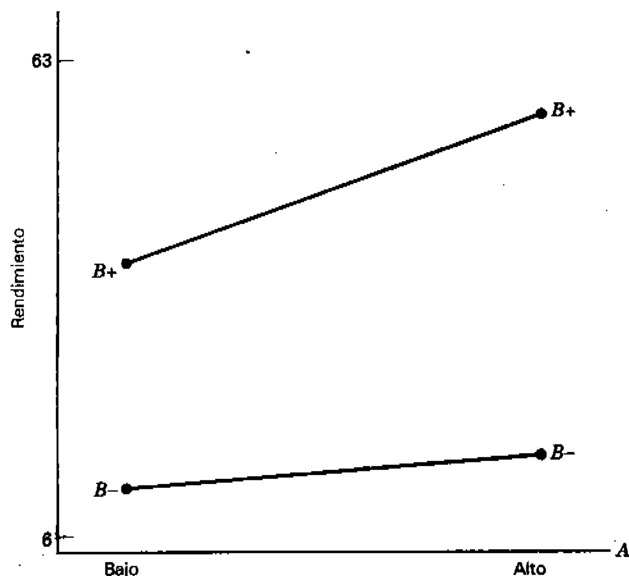


Figura 8-9 Interacción apertura-tiempo de exposición del ejemplo 8-2.

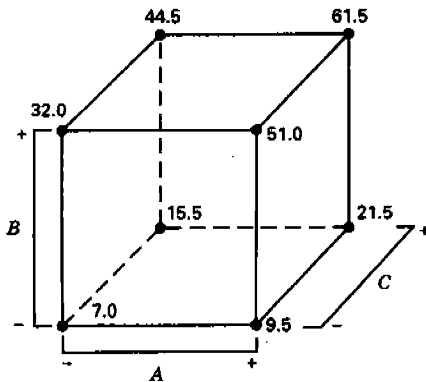


Figura 8-10 Proyección del diseño 2_V^{5-1} del ejemplo 8-2 en dos réplicas de un diseño 2^3 en los factores *A*, *B* y *C*.

El diseño 2^{5-1} se reducirá a dos réplicas de un diseño 2^3 en tres cualesquiera de los cinco factores originales. (Observar la figura 8-5 ayuda a visualizar esto.) La figura 8-10 es una gráfica de cubo en los factores *A*, *B* y *C* con los rendimientos promedio superpuestos en los ocho vértices. Es evidente por la inspección de la gráfica de cubo que los rendimientos más altos se consiguen con *A*, *B* y *C* en el nivel alto. Los factores *D* y *E* tienen un efecto pequeño sobre el rendimiento promedio del proceso y pueden ajustarse en los valores que optimicen otros objetivos (como el costo).

Secuencias de diseños factoriales fraccionados

El uso de los diseños factoriales fraccionados lleva con frecuencia a una gran economía y eficiencia en la experimentación, en particular si las corridas pueden hacerse **secuencialmente.** Por ejemplo, suponga que se están investigando $k = 4$ factores ($2^4 = 16$ corridas). Casi siempre es preferible correr un diseño fraccionado 2_{IV}^{4-1} (ocho corridas), analizar los resultados y después decidir cuál es la mejor serie de corridas que deberá correrse después. Si es necesario resolver ambigüedades, siempre puede correrse la fracción alterna y completar el diseño 2^4. Cuando se usa este método para completar el diseño, ambas fracciones un medio representan **bloques** del diseño completo con las interacciones de orden superior confundidas con los bloques (en este caso *ABCD* estaría confundida). Por lo tanto, la experimentación secuencial tiene como resultado la pérdida de información sólo en la interacción de orden más alto. Su ventaja es que en muchos casos se saca información suficiente de la fracción un medio para proceder a la siguiente etapa de la experimentación, lo cual podría implicar la incorporación o eliminación de factores, el cambio de las respuestas, o la variación de algunos de los factores en nuevos rangos. Algunas de estas posibilidades se ilustran gráficamente en la figura 8-11.

EJEMPLO 8-3

Considere nuevamente el experimento del ejemplo 8-1. Se ha usado un diseño 2_{IV}^{4-1} y se ha hecho la identificación tentativa de los tres efectos principales grandes: *A*, *C* y *D*. Hay dos efectos grandes asociados con interacciones de dos factores, $AC + BD$ y $AD + BC$. En el ejemplo 8-2 se utilizó el hecho de que el efecto principal de *B* era insignificante para concluir de manera tentativa que las interacciones importantes eran

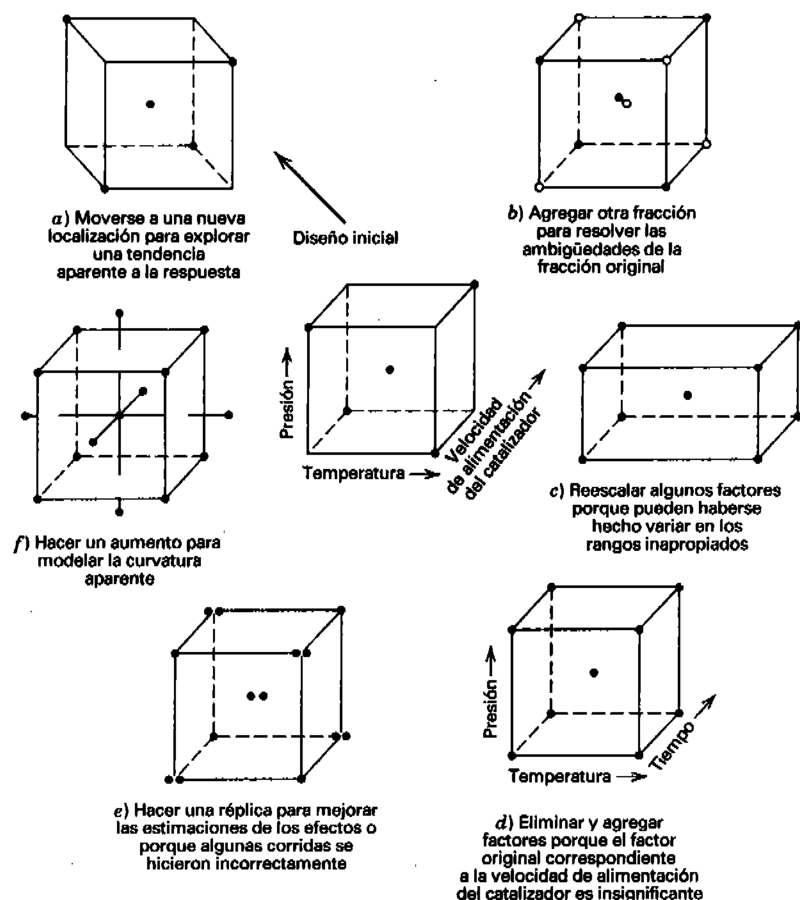


Figura 8-11 Posibilidades para el seguimiento de la experimentación después de un experimento factorial fraccionado [adaptado de Box ("Sequential Experimentation and Sequential Assembly of Designs") con permiso del editor].

AC y *AD*. En ocasiones el experimentador tendrá que procesar conocimientos que puedan ayudarle a discriminar entre las interacciones que probablemente sean importantes. Sin embargo, siempre es posible aislar la interacción significativa corriendo la fracción alterna, dada por $I = -ABCD$. Es directa la demostración de que el diseño y las respuestas son los siguientes:

	Diseño básico					
Corrida	A	B	C	$D = -ABC$	Combinación de tratamientos	Índice de filtración
1	−	−	−	+	*d*	43
2	+	−	−	−	*a*	71
3	−	+	−	−	*b*	48
4	+	+	−	+	*abd*	104
5	−	−	+	−	*c*	68
6	+	−	+	+	*acd*	86
7	−	+	+	+	*bcd*	70
8	+	+	+	−	*abc*	65

Las combinaciones lineales de las observaciones obtenidas a partir de esta fracción alterna son

$$\begin{aligned}
\ell'_A &= 24.25 \to A - BCD \\
\ell'_B &= 4.75 \to B - ACD \\
\ell'_C &= 5.75 \to C - ABD \\
\ell'_D &= 12.75 \to D - ABC \\
\ell'_{AB} &= 1.25 \to AB - CD \\
\ell'_{AC} &= -17.75 \to AC - BD \\
\ell'_{AD} &= 14.25 \to AD - BC
\end{aligned}$$

Estas estimaciones pueden combinarse con las que se obtuvieron de la fracción un medio original para obtener las siguientes estimaciones de los efectos:

i	De $\frac{1}{2}(\ell_i + \ell'_i)$	De $\frac{1}{2}(\ell_i - \ell'_i)$
A	$21.63 \to A$	$-2.63 \to BCD$
B	$3.13 \to B$	$-1.63 \to ACD$
C	$9.88 \to C$	$4.13 \to ABD$
D	$14.63 \to D$	$1.88 \to ABC$
AB	$0.13 \to AB$	$-1.13 \to CD$
AC	$-18.13 \to AC$	$-0.38 \to BD$
AD	$16.63 \to AD$	$2.38 \to BC$

Estas estimaciones concuerdan exactamente con las del análisis original de los datos como una sola réplica de un diseño factorial 2^4, como se consigna en el ejemplo 6-2. Evidentemente, son las interacciones AC y AD las que son grandes.

Agregar la fracción alterna a la fracción principal puede considerarse como un tipo de **experimento de confirmación**, por cuanto proporciona información que permitirá fortalecer las conclusiones iniciales acerca de los efectos de la interacción de dos factores. En la sección 8-5 se investigarán otros aspectos de la combinación de diseños factoriales fraccionados para aislar las interacciones. En ocasiones un experimento de confirmación no es tan elaborado como éste. Por ejemplo, podría usarse la ecuación del modelo para predecir la respuesta en un punto de interés dentro del espacio del diseño (no uno de los puntos del diseño actual), correr después realmente ese ensayo (quizá varias veces) y usar la comparación entre la respuesta predicha y la observada para confirmar los resultados.

8-3 LA FRACCIÓN UN CUARTO DEL DISEÑO 2^k

Para un número moderadamente grande de factores, con frecuencia son útiles fracciones menores del diseño 2^k. Considere una fracción un cuarto del diseño 2^k. Este diseño contiene 2^{k-2} corridas y es común llamarlo **diseño factorial fraccionado 2^{k-2}.**

El diseño 2^{k-2} puede construirse apuntando primero un **diseño básico** compuesto por las corridas asociadas con un diseño factorial completo en $k-2$ factores y asociando después las dos columnas adicionales con las interacciones elegidas apropiadamente que incluyan los primeros $k-2$ factores. Por lo tanto,

una fracción un cuarto del diseño 2^k tiene dos generadores. Si P y Q representan los generadores escogidos, entonces a $I = P$ e $I = Q$ se les llama las **relaciones generadoras** del diseño. Los signos de P y Q (+ o –) determinan cuál de las fracciones un cuarto se produce. Las cuatro fracciones asociadas con la elección de los **generadores** $\pm P$ y $\pm Q$ pertenecen a la misma **familia.** La fracción para la que tanto P como Q son positivas es la fracción principal.

La **relación de definición completa** del diseño está compuesta por todas las columnas que son iguales a la columna identidad I. Éstas constarán de P, Q y su **interacción generalizada** PQ; es decir, la relación de definición es $I = P = Q = PQ$. A los elementos P, Q y PQ de la relación de definición se les denomina **palabras.** Los alias de cualquier efecto se obtienen mediante la multiplicación de la columna de ese efecto por cada palabra de la relación de definición. Evidentemente, cada efecto tiene *tres* alias. El experimentador deberá estar atento al elegir los generadores para que los efectos potencialmente importantes no sean alias entre sí.

Como un ejemplo, considere el diseño 2^{6-2}. Suponga que se escogen $I = ABCE$ e $I = BCDF$ como los generadores del diseño. Entonces la interacción generalizada de los generadores $ABCE$ y $BCDF$ es $ADEF$; por lo tanto, la relación de definición completa de este diseño es

$$I = ABCE = BCDF = ADEF$$

Por consiguiente, se trata de un diseño de resolución IV. Para encontrar los alias de cualquier efecto (por ejemplo de A), se multiplica ese efecto por cada palabra de la relación de definición. Para A, esto produce

$$A = BCE = ABCDF = DEF$$

Es sencillo verificar que todos los efectos principales son alias de interacciones de tres y cinco factores, mientras que las interacciones de dos factores son alias entre sí y de interacciones de órdenes superiores. Por lo tanto, cuando se estima A, por ejemplo, en realidad se está estimando $A + BCE + DEF + ABCDF$. En la tabla 8-8 se muestra la estructura completa de los alias de este diseño. Si las interacciones de tres factores y de órdenes superiores son insignificantes, este diseño produce estimaciones claras de los efectos principales.

Para construir este diseño se anota primero el **diseño básico**, el cual consiste en las 16 corridas para un diseño completo $2^{6-2} = 2^4$ en A, B, C y D. Después se añaden los dos factores E y F, asociando sus niveles más y menos con los signos más y menos de la interacción ABC y BCD, respectivamente. Este procedimiento se muestra en la tabla 8-9.

Otra forma de construir este diseño es deduciendo los cuatro bloques del diseño 2^6 con $ABCE$ y $BCDF$ confundidas y eligiendo después el bloque con las combinaciones de tratamientos que son positivas para $ABCE$ y $BCDF$. Se trataría de un diseño factorial fraccionado 2^{6-2} con relaciones generadoras $I = ABCE$ e $I = BCDF$, y puesto que los dos generadores $ABCE$ y $BCDF$ son positivos, se trata de la fracción principal.

Tabla 8-8 Estructura de los alias del diseño 2^{6-2}_{IV} con $I = ABCE = BCDF = ADEF$

$A = BCE = DEF = ABCDF$	$AB = CE = ACDF = BDEF$
$B = ACE = CDF = ABDEF$	$AC = BE = ABDF = CDEF$
$C = ABE = BDF = ACDEF$	$AD = EF = BCDE = ABCF$
$D = BCF = AEF = ABCDE$	$AE = BC = DF = ABCDEF$
$E = ABC = ADF = BCDEF$	$AF = DE = BCEF = ABCD$
$F = BCD = ADE = ABCEF$	$BD = CF = ACDE = ABEF$
$ABD = CDE = ACF = BEF$	$BF = CD = ACEF = ABDE$
$ACD = BDE = ABF = CEF$	

Tabla 8-9 Construcción del diseño 2_{IV}^{6-2} con los generadores $I = ABCE$ e $I = BCDF$

	Diseño básico					
Corrida	A	B	C	D	$E = ABC$	$F = BCD$
1	−	−	−	−	−	−
2	+	−	−	−	+	−
3	−	+	−	−	+	+
4	+	+	−	−	−	+
5	−	−	+	−	+	+
6	+	−	+	−	−	+
7	−	+	+	−	−	−
8	+	+	+	−	+	−
9	−	−	−	+	−	+
10	+	−	−	+	+	+
11	−	+	−	+	+	−
12	+	+	−	+	−	−
13	−	−	+	+	+	−
14	+	−	+	+	−	−
15	−	+	+	+	−	+
16	+	+	+	+	+	+

Hay, desde luego, tres **fracciones alternas** de este diseño 2_{IV}^{6-2} particular. Se trata de las fracciones con las relaciones generadoras $I = ABCE$ e $I = -BCDF$; $I = -ABCE$ e $I = BCDF$; e $I = -ABCE$ e $I = -BCDF$. Es sencillo construir estas fracciones con el método que se muestra en la tabla 8-9. Por ejemplo, si quiere encontrarse la fracción para la que $I = ABCE$ e $I = -BCDF$, entonces en la última columna de la tabla 8-9 se hace $F = -BCD$, y la columna de los niveles del factor F queda como

$$++----++--++++--$$

La relación de definición completa de esta fracción alterna es $I = ABCE = -BCDF = -ADEF$. Ahora ciertos signos en la estructura de los alias de la tabla 8-9 se han cambiado; por ejemplo, los alias de A son $A = BCE = -DEF = -ABCDF$. Por lo tanto, la combinación lineal de las observaciones ℓ_A estima en realidad $A + BCE - DEF - ABCDF$.

Por último, observe que el diseño factorial fraccionado 2_{IV}^{6-2} se proyectará en una sola réplica de un diseño 2^4 en cualquier subconjunto de cuatro factores que no sea una palabra de la relación de definición. También se pliega en una fracción un medio con una réplica de un diseño 2^4 en cualquier subconjunto de cuatro factores que sea una palabra de la relación de definición. Por lo tanto, el diseño de la tabla 8-9 se convierte en dos réplicas de un diseño 2^{4-1} en los factores $ABCE$, $BCDF$ y $ADEF$, porque éstas son las palabras de la relación de definición. Hay otras 12 combinaciones de los seis factores, como $ABCD$, $ABCF$, etc., para las que el diseño se proyecta en una sola réplica del diseño 2^4. Este diseño también se pliega en dos réplicas de un diseño 2^3 en *cualquier* subconjunto de tres de los seis factores o en cuatro réplicas de un diseño 2^2 en cualquier subconjunto de dos factores.

En general, cualquier diseño factorial fraccionado 2^{k-2} puede plegarse en un diseño factorial completo o bien en un diseño factorial fraccionado en algún subconjunto de $r \le k - 2$ de los factores originales. Estos subconjuntos de variables que forman diseños factoriales completos no son palabras de la relación de definición completa.

EJEMPLO 8-4 ....

Las piezas fabricadas en un proceso de moldeo por inyección están presentando una contracción excesiva. Esto está ocasionando problemas en las operaciones de ensamblaje que se realizan después del moldeo por inyección. Un equipo de mejoramiento de calidad ha decidido llevar a cabo un experimento

diseñado para estudiar el proceso de moldeo por inyección a fin de poder reducir la contracción. El equipo decide investigar seis factores —la temperatura de moldeo (A), la velocidad del enroscado (B), el tiempo de retención (C), la duración del ciclo (D), el tamaño del vaciadero (E) y la presión de la retención (F)— con dos niveles cada uno, con el fin de saber cómo se afecta la contracción debido a cada factor, así como para obtener información preliminar acerca de la forma en que los factores interactúan.

El equipo decide usar el diseño factorial fraccionado de 16 corridas con dos niveles de la tabla 8-9. El diseño se muestra de nuevo en la tabla 8-10, junto con la contracción observada (× 10) en la pieza de prueba producida en cada una de las 16 corridas del diseño. En la tabla 8-11 se muestran las estimaciones de los efectos, las sumas de cuadrados y los coeficientes de regresión de este experimento.

En la figura 8-12 se presenta la gráfica de probabilidad normal de las estimaciones de los efectos de este experimento. Los únicos efectos grandes son A (temperatura de moldeo), B (velocidad del enroscado) y la interacción AB. A la luz de las relaciones de los alias de la tabla 8-8, parece razonable adoptar estas conclusiones de manera tentativa. La gráfica de la interacción AB de la figura 8-13 indica que el proceso muestra una alta insensibilidad a la temperatura si la velocidad del enroscado está en el nivel bajo, pero que es muy sensible a la temperatura si la velocidad del enroscado está en el nivel alto. Con la velocidad del enroscado en el nivel bajo, el proceso deberá producir una contracción promedio de alrededor de 10%, independientemente del nivel de temperatura elegido.

Con base en este análisis inicial, el equipo decide hacer el ajuste de la temperatura de moldeo y la velocidad del enroscado en el nivel bajo. Este conjunto de condiciones reducirá la contracción *media* de las piezas en alrededor de 10%. Sin embargo, la variabilidad de la contracción de una pieza a otra sigue siendo un problema potencial. De hecho, la contracción media puede reducirse adecuadamente mediante las modificaciones anteriores; sin embargo, la variabilidad de la contracción de una pieza a otra en una corrida de producción podría seguir causando problemas en el ensamblaje. Una manera de abordar esta cuestión es investigando si alguno de los factores del proceso afecta la *variabilidad* de la contracción de las piezas.

En la figura 8-14 se presenta la gráfica de probabilidad normal de los residuales. Esta gráfica parece ser satisfactoria. Se construyeron después las gráficas de los residuales contra cada factor. En la figura

Tabla 8-10 Un diseño 2_{IV}^{6-2} para el experimento del moldeo por inyección del ejemplo 8-4

	Diseño básico						
Corrida	A	B	C	D	$E = ABC$	$F = BCD$	Contracción observada (× 10)
1	−	−	−	−	−	−	6
2	+	−	−	−	+	−	10
3	−	+	−	−	+	+	32
4	+	+	−	−	−	+	60
5	−	−	+	−	+	+	4
6	+	−	+	−	−	+	15
7	−	+	+	−	−	−	26
8	+	+	+	−	+	−	60
9	−	−	−	+	−	+	8
10	+	−	−	+	+	+	12
11	−	+	−	+	+	−	34
12	+	+	−	+	−	−	60
13	−	−	+	+	+	−	16
14	+	−	+	+	−	−	5
15	−	+	+	+	−	+	37
16	+	+	+	+	+	+	52

Tabla 8-11 Efectos, sumas de cuadrados y coeficientes de regresión del ejemplo 8-4

Variable	Nombre	Nivel −1	Nivel +1
A	temperatura_moldeo	−1.000	1.000
B	velocidad_enroscado	−1.000	1.000
C	duración_retención	−1.000	1.000
D	duración_ciclo	−1.000	1.000
E	tamaño_vaciadero	−1.000	1.000
F	presión_retención	−1.000	1.000

Variable[a]	Coeficiente de regresión	Efecto estimado	Suma de cuadrados
Promedio global	27.3125		
A	6.9375	13.8750	770.062
B	17.8125	35.6250	5076.562
C	−0.4375	−0.8750	3.063
D	0.6875	1.3750	7.563
E	0.1875	0.3750	0.563
F	0.1875	0.3750	0.563
AB + *CE*	5.9375	11.8750	564.063
AC + *BE*	−0.8125	−1.6250	10.562
AD + *EF*	−2.6875	−5.3750	115.562
AE + *BC* + *DF*	−0.9375	−1.8750	14.063
AF + *DE*	0.3125	0.6250	1.563
BD + *CF*	−0.0625	−0.1250	0.063
BF + *CD*	−0.0625	−0.1250	0.063
ABD	0.0625	0.1250	0.063
ABF	−2.4375	−4.8750	95.063

[a]Sólo los efectos principales y las interacciones de dos factores.

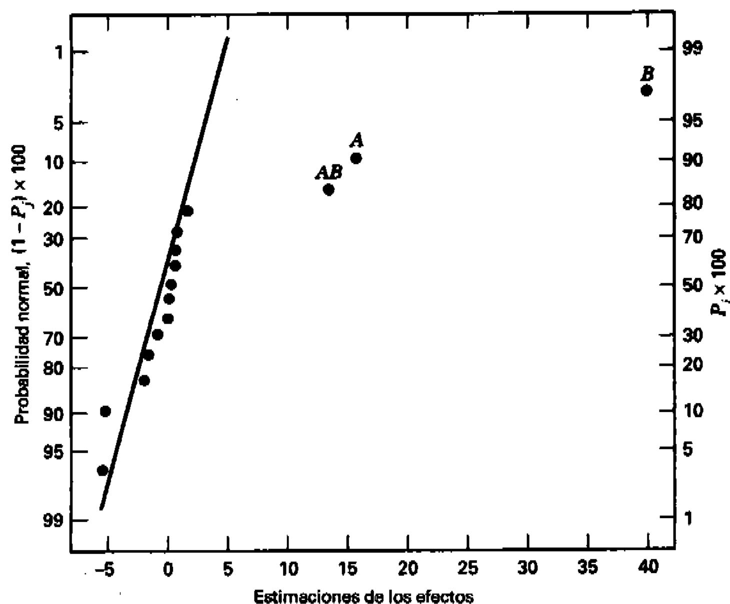


Figura 8-12 Gráfica de probabilidad normal de los efectos del ejemplo 8-4.

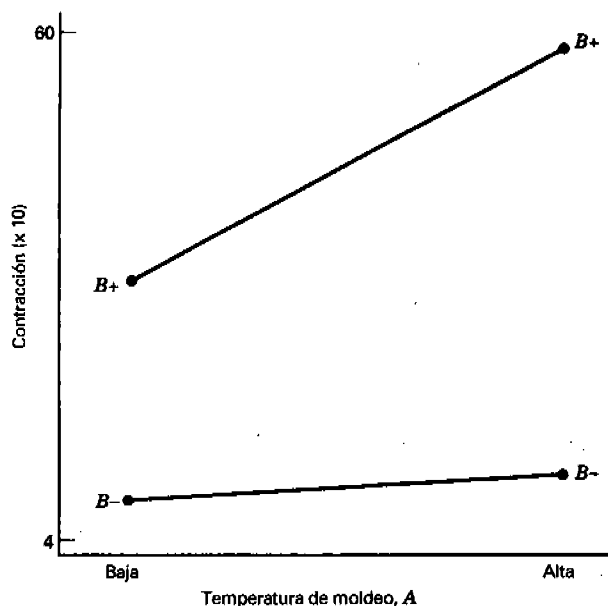


Figura 8-13 Gráfica de la interacción *AB* (temperatura de moldeo-velocidad del enroscado) para el ejemplo 8-4.

8-15 se muestra una de estas gráficas, la de los residuales contra el factor C (tiempo de retención). La gráfica revela que hay una dispersión sensiblemente menor en los residuales con el tiempo de retención bajo que con el tiempo de retención alto. Estos residuales se obtuvieron de la manera usual a partir del modelo de la contracción predicha:

$$\begin{aligned}\hat{y} &= \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \hat{\beta}_{12} x_1 x_2 \\ &= 27.3125 + 6.9375 x_1 + 17.8125 x_2 + 5.9375 x_1 x_2\end{aligned}$$

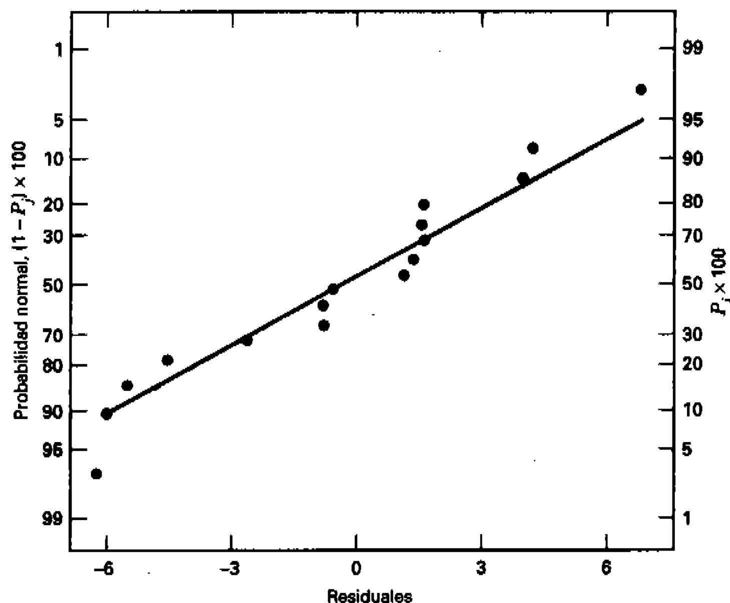


Figura 8-14 Gráfica de probabilidad normal de los residuales del ejemplo 8-4.

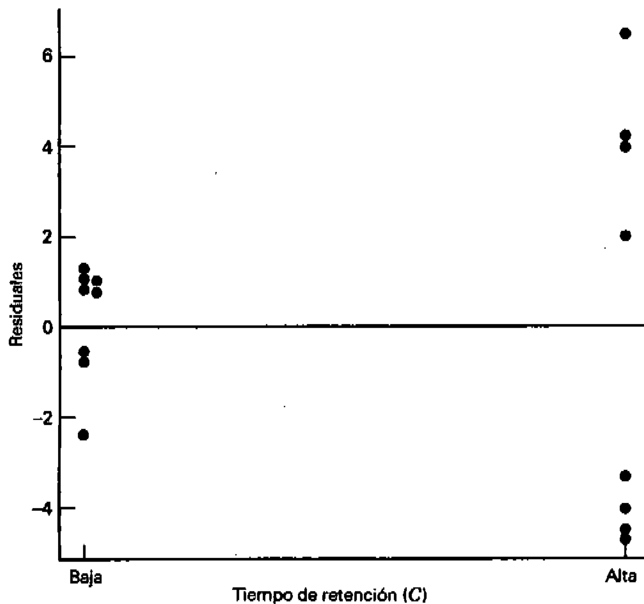


Figura 8-15 Residuales contra el tiempo de retención (C) para el ejemplo 8-4.

donde x_1, x_2 y x_1x_2 son las variables codificadas que corresponden a los factores A y B y a la interacción AB. Entonces los residuales son

$$e = y - \hat{y}$$

El modelo de regresión usado para producir los residuales elimina, en esencia, los **efectos de localización** de A, B y AB de los datos; por lo tanto, los residuales contienen información acerca de la variabilidad no explicada. La figura 8-15 indica que existe un *patrón* en la variabilidad y que la variabilidad de la contracción de las piezas puede ser menor cuando el tiempo de retención está en el nivel bajo (recuerde que en el capítulo 6 se señaló que los residuales sólo transmiten información acerca de los efectos de dispersión cuando es correcto el modelo de localización o la media).

Lo anterior se observa con mayor claridad en el análisis de los residuales que se presenta en la tabla 8-12. En esta tabla, los residuales se ordenan en los niveles bajo (–) y alto (+) de cada factor, y se ha calculado la desviación estándar de los residuales en los niveles bajo y alto de cada factor. Observe que la desviación estándar de los residuales con C en el nivel bajo [$S(C^-) = 1.63$] es considerablemente menor que la desviación estándar de los residuales con C en el nivel alto [$S(C^+) = 5.70$].

En el último renglón de la tabla 8-12 se presenta el estadístico

$$F_i^* = \ln \frac{S^2(i^+)}{S^2(i^-)}$$

Recuerde que si las varianzas de los residuales en los niveles alto (+) y bajo (–) del factor i son iguales, entonces este cociente sigue una distribución aproximadamente normal con media cero, y puede usarse para evaluar la diferencia en la variabilidad de la respuesta en los dos niveles del factor i. Puesto que el cociente F_C^* es relativamente grande, se concluiría que la aparente **dispersión** o **efecto de variabilidad** observado en la figura 8-15 es real. Por lo tanto, ajustar el tiempo de retención en su nivel bajo contribuiría a reducir la variabilidad de una pieza a otra durante una corrida de producción. En la figura 8-16 se presen-

Tabla 8-12 Cálculo de los efectos de dispersión del ejemplo 8-4

Corrida	A	B	$AB = CE$	C	$AC = BE$	$AE = BC = DF$	E	D	$AD = EF$	$BD = CE$	ABD	$BF = CD$	ACD	F	$AF = DE$	Residual
1	–	–	+	–	+	+	–	–	+	+	–	+	–	–	+	–2.50
2	+	–	–	–	–	+	+	–	–	+	+	+	+	–	–	–0.50
3	–	+	–	–	+	–	+	–	+	–	+	+	–	+	–	–0.25
4	+	+	+	–	–	–	–	–	–	–	–	+	+	+	+	2.00
5	–	–	+	+	–	–	+	–	+	+	–	–	+	+	–	–4.50
6	+	–	–	+	+	–	–	–	–	+	+	–	–	+	+	4.50
7	–	+	–	+	–	+	–	–	+	–	+	–	+	–	+	–6.25
8	+	+	+	+	+	+	+	–	–	–	–	–	–	–	–	2.00
9	–	–	+	–	+	+	–	+	–	–	+	–	+	+	–	–0.50
10	+	–	–	–	–	+	+	+	+	–	–	–	–	+	+	1.50
11	–	+	–	–	+	–	+	+	–	+	–	–	+	–	+	1.75
12	+	+	+	–	–	–	–	+	+	+	+	–	–	–	–	2.00
13	–	–	+	+	–	–	+	+	–	–	+	+	–	–	+	7.50
14	+	–	–	+	+	–	–	+	+	–	–	+	+	–	–	–5.50
15	–	+	–	+	–	+	–	+	–	+	–	+	–	+	–	4.75
16	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	–6.00
$S(i^+)$	3.80	4.01	4.33	5.70	3.68	3.85	4.17	4.64	3.39	4.01	4.72	4.71	3.50	3.88	4.87	
$S(i^-)$	4.60	4.41	4.10	1.63	4.53	4.33	4.25	3.59	2.75	4.41	3.51	3.65	3.12	4.52	3.40	
F_i^*	–0.38	–0.19	0.11	2.50	–0.42	–0.23	–0.04	0.51	0.42	–0.19	0.59	0.51	0.23	–0.31	0.72	

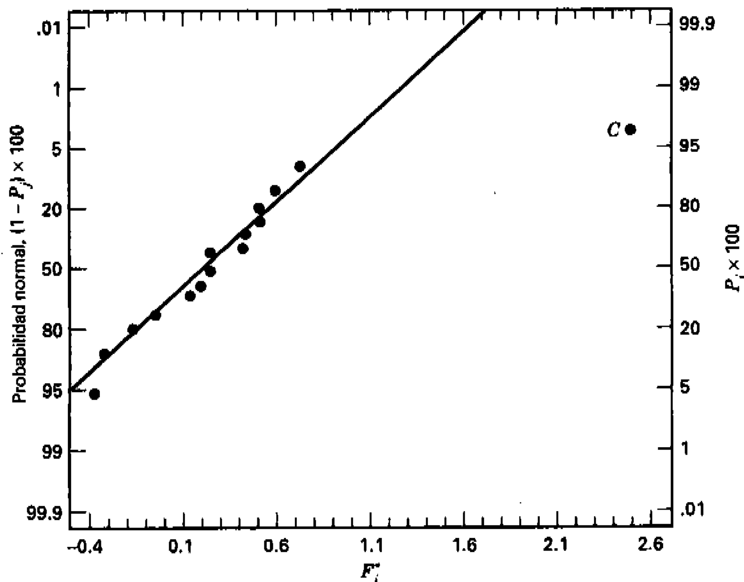


Figura 8-16 Gráfica de probabilidad normal de los efectos de dispersión F_i^* del ejemplo 8-4.

ta una gráfica de probabilidad normal de los valores F_i^* de la tabla 8-12; ésta también indica que el factor C tiene un efecto de dispersión grande.

En la figura 8-17 se muestran los datos de este experimento proyectados en un cubo en los factores A, B y C. La contracción promedio observada y el rango de la contracción observada se indican en cada vértice del cubo. Por la inspección de la figura se observa que correr el proceso con la velocidad del enroscado (B) en el nivel bajo es la clave para reducir la contracción promedio de las piezas. Si B está en el nivel bajo, virtualmente cualquier combinación de la temperatura (A) y el tiempo de retención (C) resultará en valores bajos de la contracción promedio de las piezas. Sin embargo, al examinar los rangos de los valores de la contracción en cada vértice del cubo, es claro de inmediato que ajustar el tiempo de retención (C) en el nivel bajo es la única elección razonable si se quiere mantener baja la variabilidad de la contracción de una pieza a otra en una corrida de producción.

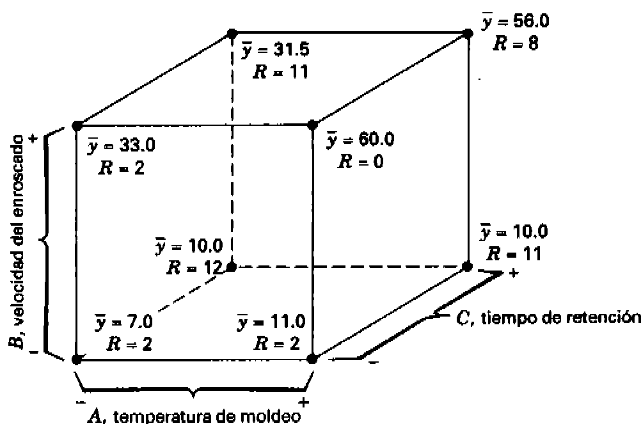


Figura 8-17 Contracción promedio y rango de la contracción en los factores A, B y C para el ejemplo 8-4.

8-4 EL DISEÑO FACTORIAL FRACCIONADO 2^{k-p} GENERAL

A un diseño factorial fraccionado 2^k que contiene 2^{k-p} corridas se le llama fracción $1/2^p$ del diseño 2^k o, de manera más simple, **diseño factorial fraccionado 2^{k-p}.** En estos diseños deben seleccionarse p generadores independientes. La relación de definición de este diseño se compone de los p generadores elegidos inicialmente y sus $2^p - p - 1$ interacciones generalizadas. En la presente sección se estudia la construcción y el análisis de estos diseños.

La estructura de los alias puede encontrarse multiplicando la columna de cada efecto por la relación de definición. Deberá prestarse atención al elegir los generadores para que los efectos de interés potencial no sean alias entre sí. Cada efecto tiene $2^p - 1$ alias. Para valores moderadamente grandes de k, es común suponer que las interacciones de órdenes superiores (por ejemplo, de tercero y cuarto orden y superiores) son insignificantes, con lo cual se simplifica en gran medida la estructura de los alias.

Es importante seleccionar los p generadores de un diseño factorial fraccionado 2^{k-p} de tal modo que se obtengan las mejores relaciones de los alias posibles. Un criterio razonable es seleccionar los generadores para que el diseño 2^{k-p} resultante tenga la **resolución más alta posible.** Para ilustrar, considere el diseño 2_{IV}^{6-2} de la tabla 8-9, donde se usaron los generadores $E = ABC$ y $F = BCD$, con lo cual se produce un diseño de resolución IV. Éste es el diseño con la resolución más alta. Si se hubieran seleccionado $E = ABC$ y $F = ABCD$, la relación de definición completa hubiera sido $I = ABCE = ABCDF = DEF$, y el diseño habría sido de resolución III. Se trata, evidentemente, de una elección inferior porque sacrifica de manera innecesaria información acerca de las interacciones.

En ocasiones la resolución por sí sola no es suficiente para distinguir entre los diseños. Por ejemplo, considere los tres diseños 2_{IV}^{7-2} de la tabla 8-13. Todos estos diseños son de resolución IV, pero tienen estructuras de los alias muy diferentes (se ha supuesto que las interacciones de tres factores y las de órdenes superiores son insignificantes) con respecto a las interacciones de dos factores. Evidentemente, el diseño A es el que tiene más alias y el diseño C el que tiene menos, por lo que el diseño C sería una buena elección para un diseño 2_{IV}^{7-2}.

Las tres palabras del diseño A tienen longitud 4; es decir, el **patrón** de la longitud de las palabras es $\{4, 4, 4\}$. Para el diseño B es $\{4, 4, 6\}$ y para el diseño C es $\{4, 5, 5\}$. Observe que la relación de definición del diseño C tiene una sola palabra de cuatro letras, mientras que los demás diseños tienen dos o tres. Por lo tanto, el diseño C minimiza el número de palabras de la relación de definición que son de longitud mínima. A un diseño como éste se le llama **diseño de aberración mínima.** Minimizar la aberración en un diseño de resolución R asegura que el diseño tiene el número mínimo de efectos principales que son alias de

Tabla 8-13 Tres elecciones de generadores para el diseño 2_{IV}^{7-2}

Generadores del diseño A: $F = ABC$, $G = BCD$, $I = ABCF = BCDG = ADFG$	Generadores del diseño B: $F = ABC$, $G = ADE$, $I = ABCF = ADEG = BCDEFG$	Generadores del diseño C: $F = ABCD$, $G = ABDE$, $I = ABCDF = ABDEG = CEFG$
Alias (interacciones de dos factores)	Alias (interacciones de dos factores)	Alias (interacciones de dos factores)
$AB = CF$	$AB = CF$	$CE = FG$
$AC = BF$	$AC = BF$	$CF = EG$
$AD = FG$	$AD = EG$	$CG = EF$
$AG = DF$	$AE = DG$	
$BD = CG$	$AF = BC$	
$BG = CD$	$AG = DE$	
$AF = BC = DG$		

interacciones de orden $R - 1$, el número mínimo de interacciones de dos factores que son alias de interacciones de orden $R - 2$, etcétera. Referirse a Fries y Hunter [46] para mayores detalles.

En la tabla 8-14 se presenta una selección de diseños factoriales fraccionados 2^{k-p} para $k \leq 15$ factores y hasta $n \leq 128$ corridas. Los generadores sugeridos en esta tabla resultarán en un diseño con la resolución más alta posible. Son también los diseños con aberración mínima.

Las relaciones de los alias para todos los diseños de la tabla 8-14 para los que $n \leq 64$ se presentan en la tabla XII(a-w) del apéndice. Las relaciones de los alias incluidas en esta tabla se enfocan en los efectos principales y las interacciones de dos y tres factores. Se da la relación de definición completa para cada diseño. Esta tabla del apéndice hace muy sencillo seleccionar un diseño con la resolución suficiente para asegurar que cualesquiera interacciones de interés potencial puedan estimarse.

EJEMPLO 8-5

Para ilustrar el uso de la tabla 8-14, suponga que se tienen siete factores y que el interés se encuentra en estimar los siete efectos principales y hacerse una idea aproximada de las interacciones de dos factores. Estamos dispuestos a suponer que las interacciones de tres factores y de órdenes superiores son insignificantes. Esta información sugiere que un diseño de resolución IV sería apropiado.

La tabla 8-14 muestra que se cuenta con dos fracciones de resolución IV: la 2_{IV}^{7-2} con 32 corridas y la 2_{IV}^{7-3} con 16 corridas. La tabla XII del apéndice contiene las relaciones de los alias completas para estos dos diseños. Los alias para el diseño 2_{IV}^{7-3} de 16 corridas se encuentran en la tabla XII(i) del apéndice. Observe que los siete efectos principales son alias de interacciones de tres factores. Las interacciones de dos factores son alias en grupos de tres. Por lo tanto, este diseño satisfará los objetivos del problema; es decir, permitirá la estimación de los efectos principales y dará cierta idea respecto de las interacciones de dos factores. No es necesario correr el diseño 2_{IV}^{7-2}, el cual requeriría 32 corridas. La tabla XII(j) del apéndice indica que este diseño permitiría la estimación de los siete efectos principales y que 15 de las 21 interacciones de dos factores también podrían estimarse de manera única. (Recuerde que las interacciones de tres factores y de órdenes superiores son insignificantes.) Ésta es más de la información necesaria acerca de las interacciones. El diseño completo se muestra en la tabla 8-15. Observe que se construyó empezando con la corrida 16 del diseño 2^4 en A, B, C y D como el diseño básico y agregando después las tres columnas $E = ABC, F = BCD$ y $G = ACD$. Los generadores son $I = ABCE, I = BCDF$ e $I = ACDG$ (tabla 8-14). La relación de definición completa es $I = ABCE = BCDF = ADEF = ACDG = BDEG = CEFG = ABFG$.

Análisis de los diseños factoriales fraccionados 2^{k-p}

Hay varios programas de computadora que pueden usarse para analizar el diseño factorial fraccionado 2^{k-p}. Por ejemplo, el programa *Design-Expert* ilustrado en el capítulo 6 tiene esta capacidad.

El diseño también puede analizarse recurriendo a los principios básicos; el efecto i-ésimo se estima con

$$\ell_i = \frac{2(\text{Contraste}_i)}{N} = \frac{\text{Contraste}_i}{(N/2)}$$

donde el Contraste$_i$ se encuentra utilizando los signos positivos y negativos de la columna i y donde $N = 2^{k-p}$ es el número total de observaciones. El diseño 2^{k-p} sólo permite la estimación de $2^{k-p} - 1$ efectos (y sus alias).

Tabla 8-14 Diseños factoriales fraccionados 2^{k-p} seleccionados

Número de factores, k	Fracción	Número de corridas	Generadores del diseño
3	2_{III}^{3-1}	4	$C = \pm AB$
4	2_{IV}^{4-1}	8	$D = \pm ABC$
5	2_{V}^{5-1}	16	$E = \pm ABCD$
	2_{III}^{5-2}	8	$D = \pm AB$
			$E = \pm AC$
6	2_{VI}^{6-1}	32	$F = \pm ABCDE$
	2_{IV}^{6-2}	16	$E = \pm ABC$
			$F = \pm BCD$
	2_{III}^{6-3}	8	$D = \pm AB$
			$E = \pm AC$
			$F = \pm BC$
7	2_{VII}^{7-1}	64	$G = \pm ABCDEF$
	2_{IV}^{7-2}	32	$F = \pm ABCD$
			$G = \pm ABDE$
	2_{III}^{7-3}	16	$E = \pm ABC$
			$F = \pm BCD$
			$G = \pm ACD$
	2_{III}^{7-4}	8	$D = \pm AB$
			$E = \pm AC$
			$F = \pm BC$
			$G = \pm ABC$
8	2_{V}^{8-2}	64	$G = \pm ABCD$
			$H = \pm ABEF$
	2_{IV}^{8-3}	32	$F = \pm ABC$
			$G = \pm ABD$
			$H = \pm BCDE$
	2_{IV}^{8-4}	16	$E = \pm BCD$
			$F = \pm ACD$
			$G = \pm ABC$
			$H = \pm ABD$
9	2_{VI}^{9-2}	128	$H = \pm ACDFG$
			$J = \pm BCEFG$
	2_{IV}^{9-3}	64	$G = \pm ABCD$
			$H = \pm ACEF$
			$J = \pm CDEF$
	2_{IV}^{9-4}	32	$F = \pm BCDE$
			$G = \pm ACDE$
			$H = \pm ABDE$
			$J = \pm ABCE$
	2_{III}^{9-5}	16	$E = \pm ABC$
			$F = \pm BCD$
			$G = \pm ACD$
			$H = \pm ABD$
			$J = \pm ABCD$
10	2_{V}^{10-3}	128	$H = \pm ABCG$
			$J = \pm ACDE$
			$K = \pm ACDF$

Tabla 8-14 (continuación)

Número de factores, k	Fracción	Número de corridas	Generadores del diseño
	2_{IV}^{10-4}	64	$G = \pm BCDF$
			$H = \pm ACDF$
			$J = \pm ABDE$
			$K = \pm ABCE$
	2_{IV}^{10-5}	32	$F = \pm ABCD$
			$G = \pm ABCE$
			$H = \pm ABDE$
			$J = \pm ACDE$
			$K = \pm BCDE$
	2_{III}^{10-6}	16	$E = \pm ABC$
			$F = \pm BCD$
			$G = \pm ACD$
			$H = \pm ABD$
			$J = \pm ABCD$
			$K = \pm AB$
11	2_{IV}^{11-5}	64	$G = \pm CDE$
			$H = \pm ABCD$
			$J = \pm ABF$
			$K = \pm BDEF$
			$L = \pm ADEF$
	2_{IV}^{11-6}	32	$F = \pm ABC$
			$G = \pm BCD$
			$H = \pm CDE$
			$J = \pm ACD$
			$K = \pm ADE$
			$L = \pm BDE$
	2_{III}^{11-7}	16	$E = \pm ABC$
			$F = \pm BCD$
			$G = \pm ACD$
			$H = \pm ABD$
			$J = \pm ABCD$
			$K = \pm AB$
			$L = \pm AC$
12	2_{III}^{12-8}	16	$E = \pm ABC$
			$F = \pm ABD$
			$G = \pm ACD$
			$H = \pm BCD$
			$J = \pm ABCD$
			$K = \pm AB$
			$L = \pm AC$
			$M = \pm AD$
13	2_{III}^{13-9}	16	$E = \pm ABC$
			$F = \pm ABD$
			$G = \pm ACD$
			$H = \pm BCD$
			$J = \pm ABCD$
			$K = \pm AB$

Tabla 8-14 (continuación)

Número de factores, k	Fracción	Número de corridas	Generadores del diseño
			$L = \pm AC$
			$M = \pm AD$
			$N = \pm BC$
14	2_{III}^{14-10}	16	$E = \pm ABC$
			$F = \pm ABD$
			$G = \pm ACD$
			$H = \pm BCD$
			$J = \pm ABCD$
			$K = \pm AB$
			$L = \pm AC$
			$M = \pm AD$
			$N = \pm BC$
			$O = \pm BD$
15	2_{III}^{15-11}	16	$E = \pm ABC$
			$F = \pm ABD$
			$G = \pm ACD$
			$H = \pm BCD$
			$J = \pm ABCD$
			$K = \pm AB$
			$L = \pm AC$
			$M = \pm AD$
			$N = \pm BC$
			$O = \pm BD$
			$P = \pm CD$

Tabla 8-15 Un diseño factorial fraccionado 2_{IV}^{7-3}

	Diseño básico						
Corrida	A	B	C	D	$E = ABC$	$F = BCD$	$G = ACD$
1	−	−	−	−	−	−	−
2	+	−	−	−	+	−	+
3	−	+	−	−	+	+	−
4	+	+	−	−	−	+	+
5	−	−	+	−	+	+	+
6	+	−	+	−	−	+	−
7	−	+	+	−	−	−	+
8	+	+	+	−	+	−	−
9	−	−	−	+	−	+	+
10	+	−	−	+	+	+	−
11	−	+	−	+	+	−	+
12	+	+	−	+	−	−	−
13	−	−	+	+	+	−	−
14	+	−	+	+	−	−	+
15	−	+	+	+	−	+	−
16	+	+	+	+	+	+	+

Proyección del diseño factorial fraccionado 2^{k-p}

El diseño 2^{k-p} se reduce a un factorial completo o bien a un factorial fraccionado en cualquier subconjunto de $r \leq k-p$ de los factores originales. Esos subconjuntos de factores que producen diseños factoriales fraccionados son subconjuntos que aparecen como palabras en la relación de definición completa. Esto resulta de particular utilidad en los experimentos de tamizado cuando se sospecha desde el principio del experimento que la mayoría de los factores originales tendrán efectos pequeños. El diseño factorial fraccionado 2^{k-p} puede proyectarse entonces en un factorial completo, por ejemplo, en los factores de mayor interés. Las conclusiones a que se llegue con diseños de este tipo deberán considerarse tentativas y someterse a análisis adicional. Por lo general es posible encontrar explicaciones alternativas de los datos que intervienen en interacciones de órdenes superiores.

Como un ejemplo, considere el diseño 2_{IV}^{7-3} del ejemplo 8-5. Se trata de un diseño con 16 corridas en el que intervienen siete factores. Se proyectará en un factorial completo en cuatro cualesquiera de los siete factores originales que no sean una palabra de la relación de definición. Hay 35 subconjuntos de cuatro factores, siete de los cuales aparecen en la relación de definición completa (ver la tabla 8-15). Por lo tanto, hay 28 subconjuntos de cuatro factores que formarían diseños 2^4. Una combinación que es obvia al inspeccionar la tabla 8-15 es *A, B, C* y *D*.

Para ilustrar apropiadamente la utilidad de esta proyección, suponga que se realiza un experimento para mejorar la eficiencia de un molino de bolas y los siete factores son los siguientes:

1. Velocidad del motor
2. Muesca
3. Modo de alimentación
4. Tamaño de la alimentación
5. Tipo de material
6. Ángulo de la criba
7. Nivel de vibración de la criba

Se tiene una certeza razonable de que la velocidad del motor, el tamaño de la alimentación y el tipo de material afectarán la eficiencia y que además estos factores pueden interactuar. Se sabe menos del papel de los otros tres factores, pero es probable que sean insignificantes. Una estrategia razonable sería asignar la velocidad del motor, el modo de alimentación, el tamaño de la alimentación y el tipo de material a las columnas *A, B, C* y *D*, respectivamente, de la tabla 8-15. La muesca, el ángulo de la criba y el nivel de vibración de la criba se asignarían a las columnas *E, F* y *G*, respectivamente. Si se está en lo correcto y las "variables menores" *E, F* y *G* son insignificantes, quedará un diseño 2^4 completo en las variables clave del proceso.

Separación en bloques de diseños factoriales fraccionados

Ocasionalmente, un diseño factorial fraccionado requiere tantas corridas que no es posible realizarlas todas bajo condiciones homogéneas. En estas situaciones, los diseños factoriales fraccionados pueden confundirse o mezclarse en bloques. La tabla XII del apéndice contiene los arreglos recomendados para la separación en bloques de varios de los diseños factoriales fraccionados de la tabla 8-14. El tamaño mínimo de los bloques para estos diseños es de ocho corridas.

Para ilustrar el procedimiento general, considere el diseño factorial fraccionado 2_{IV}^{6-2} con la relación de definición $I = ABCE = BCDF = ADEF$ que se muestra en la tabla 8-10. Este diseño fraccionado contiene 16 combinaciones de tratamientos. Suponga que quiere correrse este diseño en dos bloques con ocho combinaciones de tratamientos cada uno. Al seleccionar una interacción para confundirla con los bloques, se observa por el examen de la estructura de los alias de la tabla XII(f) del apéndice que hay dos

Bloque 1	Bloque 2
(1)	*ae*
abf	*acf*
cef	*bef*
abce	*bc*
abef	*df*
bde	*abd*
acd	*cde*
bcdf	*abcdef*

Figura 8-18 El diseño 2_{IV}^{6-2} en dos bloques con *ABD* confundida.

series de alias que incluyen únicamente interacciones de tres factores. La tabla sugiere seleccionar *ABD* (y sus alias) para confundirla con los bloques. Se obtendrían así los dos bloques que se muestran en la figura 8-18. Observe que el bloque principal contiene las combinaciones de tratamientos que tienen un número igual de letras en común con *ABD*. Son también las combinaciones de tratamientos para las que $L = x_1 + x_2 + x_4 = 0 \pmod 2$.

EJEMPLO 8-6 ··

Se usa una máquina CNC de cinco ejes para maquinar un propulsor utilizado en un motor de turbina. Los perfiles de los álabes son una característica importante de la calidad. Específicamente, es de interés la desviación del perfil del álabe del perfil especificado en el plano de ingeniería. Se corre un experimento para determinar cuáles son los parámetros de la máquina que afectan la desviación del perfil. Los ocho factores seleccionados en el diseño son los siguientes:

Factor	Nivel bajo (−)	Nivel alto (+)
A = desviación en el eje x (0.001 pulg)	0	15
B = desviación en el eje y (0.001 pulg)	0	15
C = desviación en el eje z (0.001 pulg)	0	15
D = fabricante de la herramienta	1	2
E = desviación del eje a (0.001 grados)	0	30
F = velocidad del areómetro (%)	90	110
G = altura de la plantilla sujetadora (0.001 pulg)	0	15
H = velocidad de alimentación (%)	90	110

Se selecciona un álabe de prueba en cada pieza para inspeccionarlo. La desviación del perfil se mide utilizando una máquina de medición coordenada, y la desviación estándar de la diferencia entre el perfil real y el perfil especificado se usa como la variable de respuesta.

La máquina tiene cuatro areómetros. Puesto que puede haber diferencias en los areómetros, los ingenieros del proceso piensan que éstos deberán tratarse como bloques.

Los ingenieros se sienten confiados de que las interacciones de tres o más factores no son muy importantes, pero están renuentes a ignorar las interacciones de dos factores. Por la tabla 8-14, inicialmente dos

diseños parecen ser apropiados: el diseño 2_{IV}^{8-4} con 16 corridas y el diseño 2_{IV}^{8-3} con 32 corridas. La tabla XII(l) del apéndice indica que si se usa el diseño con 16 corridas, habrá un número considerable de alias con interacciones de dos factores. Además, este diseño no puede correrse en cuatro bloques sin confundir cuatro interacciones de dos factores con los bloques. Por lo tanto, los experimentadores deciden usar el diseño 2_{IV}^{8-3} en cuatro bloques. En este diseño se confunden con los bloques una cadena de alias de interacciones de tres factores y una interacción de dos factores (*EH*) y sus alias de interacciones de tres factores. La interacción *EH* es la interacción entre la desviación del eje *a* y la velocidad de alimentación, y los ingenieros consideran que una interacción entre estas dos variables es altamente improbable.

La tabla 8-16 contiene el diseño y las respuestas resultantes en términos de desviación estándar × 10^3 pulg. Puesto que la variable de respuesta es una desviación estándar, con frecuencia es mejor efectuar el análisis después de una transformación logarítmica. En la tabla 8-17 se muestran las estimaciones de los efectos. La figura 8-19 es una gráfica de probabilidad normal de las estimaciones de los efectos, utilizando ln (desviación estándar × 10^3) como la variable de respuesta. Los únicos efectos grandes son A = desviación del eje x, B = desviación del eje y, y la cadena de alias que incluye $AD + BG$. Ahora bien, AD es la in-

Tabla 8-16 El diseño 2^{8-3} en cuatro bloques del ejemplo 8-6

	Diseño básico										
Corrida	*A*	*B*	*C*	*D*	*E*	$F = ABC$	$G = ABD$	$H = BCDE$	Bloque	Orden real de las corridas	Desviación estándar (× 10^3 pulg)
1	–	–	–	–	–	–	–	+	3	18	2.76
2	+	–	–	–	–	+	+	+	2	16	6.18
3	–	+	–	–	–	+	+	–	4	29	2.43
4	+	+	–	–	–	–	–	–	1	4	4.01
5	–	–	+	–	–	+	–	–	1	6	2.48
6	+	–	+	–	–	–	+	–	4	26	5.91
7	–	+	+	–	–	–	+	+	2	14	2.39
8	+	+	+	–	–	+	–	+	3	22	3.35
9	–	–	–	+	–	–	+	–	1	8	4.40
10	+	–	–	+	–	+	–	–	4	32	4.10
11	–	+	–	+	–	+	–	+	2	15	3.22
12	+	+	–	+	–	–	+	+	3	19	3.78
13	–	–	+	+	–	+	+	+	3	24	5.32
14	+	–	+	+	–	–	–	+	2	11	3.87
15	–	+	+	+	–	–	–	–	4	27	3.03
16	+	+	+	+	–	+	+	–	1	3	2.95
17	–	–	–	–	+	–	–	–	2	10	2.64
18	+	–	–	–	+	+	+	–	3	21	5.50
19	–	+	–	–	+	+	+	+	1	7	2.24
20	+	+	–	–	+	–	–	+	4	28	4.28
21	–	–	+	–	+	+	–	+	4	30	2.57
22	+	–	+	–	+	–	+	+	1	2	5.37
23	–	+	+	–	+	–	+	–	3	17	2.11
24	+	+	+	–	+	+	–	–	2	13	4.18
25	–	–	–	+	+	–	+	+	4	25	3.96
26	+	–	–	+	+	+	–	+	1	1	3.27
27	–	+	–	+	+	+	–	–	3	23	3.41
28	+	+	–	+	+	–	+	–	2	12	4.30
29	–	–	+	+	+	+	+	–	2	9	4.44
30	+	–	+	+	+	–	–	–	3	20	3.65
31	–	+	+	+	+	–	–	+	1	5	4.41
32	+	+	+	+	+	+	+	+	4	31	3.40

Tabla 8-17 Estimaciones de los efectos, coeficientes de regresión y sumas de cuadrados del ejemplo 8-6

Variable	Nombre	Nivel −1	Nivel +1
A	Desviación del eje *x*	0	15
B	Desviación del eje *y*	0	15
C	Desviación del eje *z*	0	15
D	Fabricante de la herramienta	1	2
E	Desviación del eje *a*	0	30
F	Velocidad del areómetro	90	110
G	Altura de la plantilla sujetadora	0	15
H	Velocidad de alimentación	90	110

Variable[a]	Coeficiente de regresión	Efecto estimado	Suma de cuadrados
Promedio global	1.28007		
A	0.14513	0.29026	0.674020
B	−0.10027	−0.20054	0.321729
C	−0.01288	−0.02576	0.005310
D	0.05407	0.10813	0.093540
E	−2.531E-04	−5.063E-04	2.050E-06
F	−0.01936	−0.03871	0.011988
G	0.05804	0.11608	0.107799
H	0.00708	0.01417	0.001606
AB + *CF* + *DG*	−0.00294	−0.00588	2.767E-04
AC + *BF*	−0.03103	−0.06206	0.030815
AD + *BG*	−0.18706	−0.37412	1.119705
AE	0.00402	0.00804	5.170E-04
AF + *BC*	−0.02251	−0.04502	0.016214
AG + *BD*	0.02644	0.05288	0.022370
AH	−0.02521	−0.05042	0.020339
BE	0.04925	0.09851	0.077627
BH	0.00654	0.01309	0.001371
CD + *FG*	0.01726	0.03452	0.009535
CE	0.01991	0.03982	0.012685
CG + *DF*	−0.00733	−0.01467	0.001721
CH	0.03040	0.06080	0.029568
DE	0.00854	0.01708	0.002334
DH	0.00784	0.01569	0.001969
EF	−0.00904	−0.01808	0.002616
EG	−0.02685	−0.05371	0.023078
EH	−0.01767	−0.03534	0.009993
FH	−0.01404	−0.02808	0.006308
GH	0.00245	0.00489	1.914E-04
ABE	0.01665	0.03331	0.008874
ABH	−0.00631	−0.01261	0.001273
ACD	−0.02717	−0.05433	0.023617

[a]Sólo los efectos principales y las interacciones de dos factores.

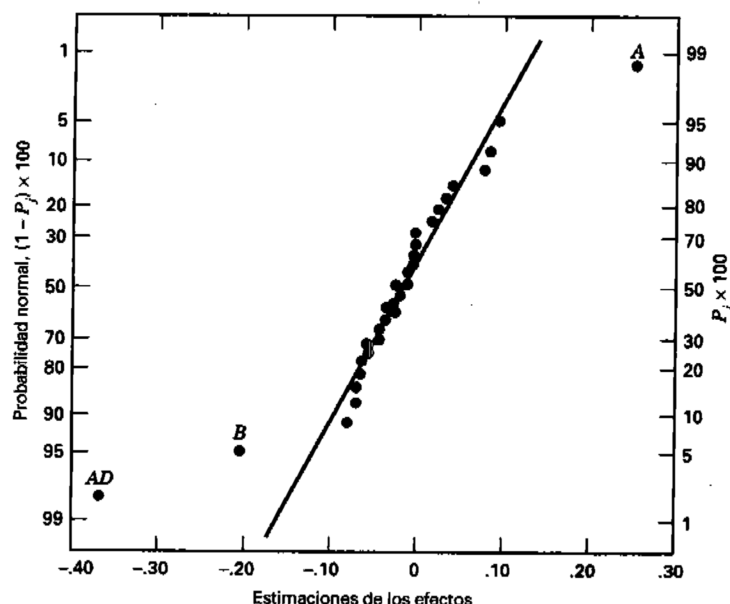


Figura 8-19 Gráfica de probabilidad normal de las estimaciones de los efectos del ejemplo 8-6.

teracción desviación del eje x-fabricante de la herramienta, y BG es la interacción desviación del eje y-altura de la plantilla sujetadora, y como estas dos interacciones son alias es imposible separarlas con base en los datos del experimento en curso. Puesto que ambas interacciones incluyen un efecto principal grande, también es difícil aplicar cualquier simplificación lógica "obvia" en esta situación. Si se contara con algún conocimiento de ingeniería o del proceso que arrojara luz sobre la situación, entonces quizá podría hacerse una elección entre las dos interacciones; en caso contrario, se necesitarán más datos para separar estos dos efectos (el problema de agregar corridas en un diseño factorial fraccionado para separar los alias de las interacciones, se estudia en la sección 8-5 y en el material suplementario de este capítulo).

Suponga que el conocimiento del proceso sugiere que posiblemente la interacción apropiada sea AD. La tabla 8-18 es el análisis de varianza resultante para el modelo con los factores A, B, D y AD (el factor D se incluyó para preservar el principio de jerarquía). Observe que el efecto del bloque es pequeño, lo cual sugiere que los areómetros de la máquina no son muy diferentes.

La figura 8-20 es una gráfica de probabilidad normal de los residuales de este experimento. Esta gráfica sugiere la presencia de colas ligeramente más gruesas que las normales, por lo que posiblemente de-

Tabla 8-18 Análisis de varianza del ejemplo 8-6

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0	Valor P
A	0.6740	1	0.6740	39.42	<0.0001
B	0.3217	1	0.3217	18.81	0.0002
D	0.0935	1	0.0935	5.47	0.0280
AD	1.1197	1	1.1197	65.48	<0.0001
Bloques	0.0201	3	0.0067		
Error	0.4099	24	0.0171		
Total	2.6389	31			

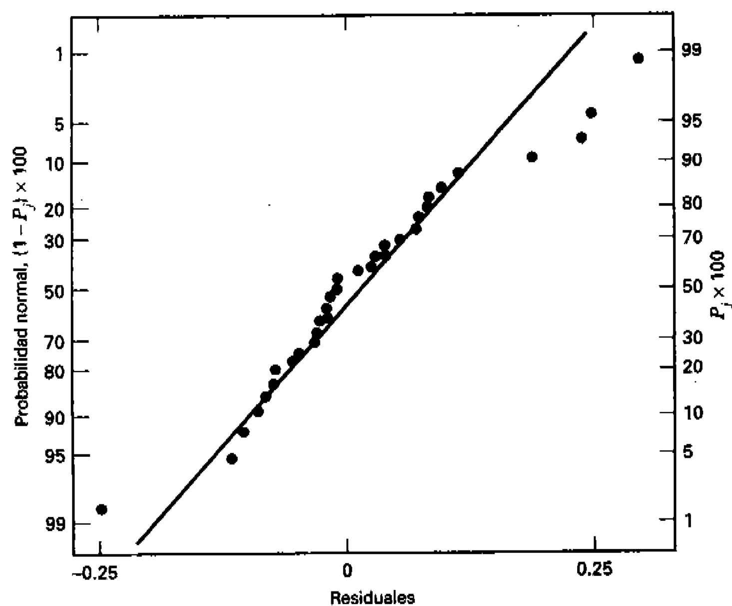


Figura 8-20 Gráfica de probabilidad normal de los residuales del ejemplo 8-6.

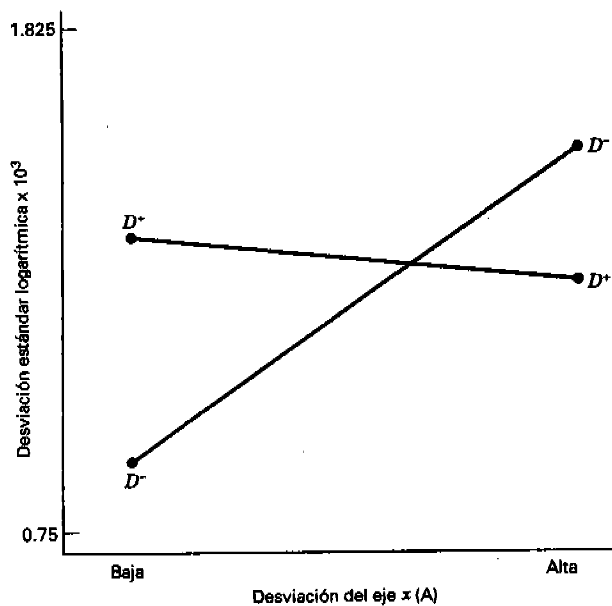


Figura 8-21 Gráfica de la interacción AD para el ejemplo 8-6.

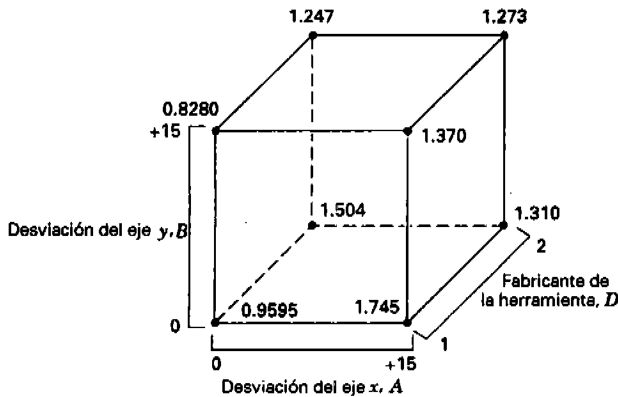


Figura 8-22 El diseño 2_{IV}^{8-3} del ejemplo 8-6 proyectado en cuatro réplicas de un diseño 2^3 en los factores A, B y D.

ban considerarse otras transformaciones. La gráfica de la interacción AD se presenta en la figura 8-21. Observe que el fabricante de la herramienta (D) y la magnitud de la desviación del eje x (A) tienen un impacto profundo en la variabilidad, del perfil del álabe, de las especificaciones de diseño. Correr A en el nivel bajo (0 desviación) y comprar las herramientas al fabricante 1 produce los mejores resultados. En la figura 8-22 se muestra la proyección de este diseño 2_{IV}^{8-3} en cuatro réplicas de un diseño 2^3 en los factores A, B y D. La mejor combinación de las condiciones de operación es A en el nivel bajo (0 desviación), B en el nivel alto (0.015 en desviación) y D en el nivel bajo (fabricante de herramientas 1).

8-5 DISEÑOS DE RESOLUCIÓN III

Como se señaló anteriormente, el uso secuencial de los diseños factoriales fraccionados es muy útil, llevando con frecuencia a una gran economía y eficiencia de la experimentación. Se ilustran ahora estas ideas utilizando la clase de los diseños de resolución III.

Es posible construir diseños de resolución III para investigar hasta $k = N - 1$ factores en sólo N corridas, donde N es un múltiplo de 4. Con frecuencia estos diseños son útiles en la experimentación industrial. Los diseños en los que N es una potencia de 2 pueden construirse con los métodos presentados anteriormente en este capítulo, y éstos se presentan primero. De particular importancia son los diseños que requieren 4 corridas para hasta 3 factores, 8 corridas para hasta 7 factores y 16 corridas para hasta 15 factores. Si $k = N - 1$, se dice que el diseño factorial fraccionado está **saturado.**

Un diseño para analizar hasta tres factores en cuatro corridas es el diseño 2_{III}^{3-1}, el cual se presentó en la sección 8-2. Otro diseño factorial fraccionado saturado muy útil es el diseño para estudiar siete factores en ocho corridas, es decir, el diseño 2_{III}^{7-4}. Este diseño es una fracción un dieciseisavo del diseño 2^7. Puede construirse apuntando primero los niveles positivos y negativos de un diseño 2^3 completo en A, B y C como el diseño básico, y asociando después los niveles de cuatro factores adicionales con las interacciones de los tres factores originales de la siguiente manera: $D = AB$, $E = AC$, $F = BC$ y $G = ABC$. Por lo tanto, los generadores de este diseño son $I = ABD$, $I = ACE$, $I = BCF$ e $I = ABCG$. El diseño se muestra en la tabla 8-19.

Tabla 8-19 El diseño 2_{III}^{7-4} con los generadores $I = ABD$, $I = ACE$, $I = BCF$ e $I = ABCG$

Corrida	Diseño básico A	B	C	$D = AB$	$E = AC$	$F = BC$	$G = ABC$	
1	–	–	–	+	+	+	–	*def*
2	+	–	–	–	–	+	+	*afg*
3	–	+	–	–	+	–	+	*beg*
4	+	+	–	+	–	–	–	*abd*
5	–	–	+	+	–	–	+	*cdg*
6	+	–	+	–	+	–	–	*ace*
7	–	+	+	–	–	+	–	*bcf*
8	+	+	+	+	+	+	+	*abcdefg*

La relación de definición completa de este diseño se obtiene multiplicando entre sí los cuatro generadores *ABD, ACE, BCF* y *ABCG* de dos en dos, de tres en tres y los cuatro a la vez, de donde se obtiene

$$I = ABD = ACE = BCF = ABCG = BCDE = ACDF = CDG$$
$$= ABEF = BEG = AFG = DEF = ADEG = CEFG = BDFG = ABCDEFG$$

Para encontrar los alias de cualquier efecto, simplemente se multiplica el efecto por cada palabra de la relación de definición. Por ejemplo, los alias de B son

$$B = AD = ABCE = CF = ACG = CDE = ABCDF = BCDG = AEF = EG$$
$$= ABFG = BDEF = ABDEG = BCEFG = DFG = ACDEFG$$

Este diseño es una fracción un dieciseisavo, y como los signos elegidos para los generadores son positivos, se trata de la fracción principal. Es también de resolución III porque el número menor de letras de cualquier palabra de la definición de contraste es tres. Cualquiera de los 16 diferentes diseños 2_{III}^{7-4} de esta familia podría construirse utilizando los generadores con 1 de los 16 arreglos posibles de los signos en $I = \pm ABD$, $I = \pm ACE$, $I = \pm BCF$, $I = \pm ABCG$.

Los siete grados de libertad de este diseño pueden usarse para estimar los siete efectos principales. Cada uno de estos efectos tiene 15 alias; sin embargo, si se supone que las interacciones de tres o más factores son insignificantes, se consigue entonces una simplificación considerable en la estructura de los alias. Estableciendo este supuesto, cada una de las combinaciones lineales asociadas con los siete efectos principales de este diseño es en realidad una estimación del efecto principal y las tres interacciones de dos factores:

$$\begin{aligned}
\ell_A &\rightarrow A + BD + CE + FG \\
\ell_B &\rightarrow B + AD + CF + EG \\
\ell_C &\rightarrow C + AE + BF + DG \\
\ell_D &\rightarrow D + AB + CG + EF \\
\ell_E &\rightarrow E + AC + BG + DF \\
\ell_F &\rightarrow F + BC + AG + DE \\
\ell_G &\rightarrow G + CD + BE + AF
\end{aligned} \tag{8-1}$$

Estos alias se encuentran en la tabla XII(h) del apéndice, ignorando las interacciones de tres factores y de órdenes superiores.

El diseño saturado 2_{III}^{7-4} de la tabla 8-19 puede usarse para obtener diseños de resolución III para estudiar menos de siete factores en ocho corridas. Por ejemplo, para generar un diseño para seis factores en ocho corridas, simplemente se elimina cualquiera de las columnas de la tabla 8-19, digamos la G. Se obtiene así el diseño que se muestra en la tabla 8-20.

Es sencillo verificar que este diseño es también de resolución III; de hecho es un diseño 2_{III}^{6-3}, o una fracción un octavo, del diseño 2^6. La relación de definición del diseño 2_{III}^{6-3} es igual a la relación de definición del diseño 2_{III}^{7-4} original, con las palabras que incluyen la letra G eliminadas. Por lo tanto, la relación de definición del nuevo diseño es

$$I = ABD = ACE = BCF = BCDE = ACDF = ABEF = DEF$$

En general, cuando se eliminan d factores para producir un nuevo diseño, la nueva relación de definición se obtiene de las palabras de la relación de definición original que no contienen ninguna de las letras eliminadas. Cuando se construyen diseños con este método, deberá prestarse atención para obtener el mejor arreglo posible. Si se eliminan las columnas B, D, F y G de la tabla 8-19, se obtiene un diseño para tres factores en ocho corridas, no obstante que las combinaciones de tratamientos corresponden a dos réplicas de un diseño 2^3. Probablemente el experimentador preferiría correr un diseño 2^3 completo en A, C y E.

También es posible obtener un diseño de resolución III para estudiar hasta 15 factores en 16 corridas. Este diseño 2_{III}^{15-11} saturado puede generarse apuntando primero las 16 combinaciones de tratamientos asociadas con un diseño 2^4 en A, B, C y D e igualando después 11 nuevos factores con las interacciones de dos, tres y cuatro factores de los cuatro factores originales. En este diseño, cada uno de los 15 efectos principales es alias de siete interacciones de dos factores. Puede usarse un procedimiento similar para el diseño 2_{III}^{31-26}, lo cual permite el estudio de hasta 31 factores en 32 corridas.

Ensamblaje secuencial de fracciones para separar efectos

Mediante la combinación de diseños factoriales fraccionados en los que se han intercambiado ciertos signos, es posible aislar de manera sistemática los efectos de interés potencial. A este tipo de experimento secuencial se le llama **doblez** o **plegado** ***(fold over)*** del diseño original. La estructura de los alias de cualquier fracción con los signos de uno o más de los factores invertidos se obtiene haciendo el cambio de signo apropiado en los factores de la estructura de los alias de la fracción original.

Tabla 8-20 El diseño 2_{III}^{6-3} con los generadores $I = ABD$, $I = ACE$ e $I = BCF$

	Diseño básico						
Corrida	A	B	C	$D = AB$	$E = AC$	$F = BC$	
1	−	−	−	+	+	+	*def*
2	+	−	−	−	−	+	*af*
3	−	+	−	−	+	−	*be*
4	+	+	−	+	−	−	*abd*
5	−	−	+	+	−	−	*cd*
6	+	−	+	−	+	−	*ace*
7	−	+	+	−	−	+	*bcf*
8	+	+	+	+	+	+	*abcdef*

Considere el diseño 2_{III}^{7-4} de la tabla 8-19. Suponga que junto con esta fracción principal se corre también un segundo diseño fraccionado con los signos invertidos en la columna del factor D. Es decir, la columna de D de la segunda fracción es

$$- + + - - + + -$$

Los efectos que pueden estimarse a partir de la primera fracción se muestran en la ecuación 8-1, y a partir de la segunda fracción se obtiene

es decir,
$$\begin{aligned}
\ell'_A &\rightarrow A - BD + CE + FG \\
\ell'_B &\rightarrow B - AD + CF + EG \\
\ell'_C &\rightarrow C + AE + BF - DG \\
\ell'_D &\rightarrow D - AB - CG - EF \\
\ell'_{-D} &\rightarrow -D + AB + CG + EF \\
\ell'_E &\rightarrow E + AC + BG - DF \\
\ell'_F &\rightarrow F + BC + AG - DE \\
\ell'_G &\rightarrow G - CD + BE + AF
\end{aligned} \tag{8-2}$$

suponiendo que no son significativas las interacciones de tres factores y de órdenes superiores. Ahora bien, a partir de las dos combinaciones lineales de los efectos $\frac{1}{2}(\ell_i + \ell'_i)$ y $\frac{1}{2}(\ell_i - \ell'_i)$ se obtiene

i	De $\frac{1}{2}(\ell_i + \ell'_i)$	De $\frac{1}{2}(\ell_i - \ell'_i)$
A	$A + CE + FG$	BD
B	$B + CF + EG$	AD
C	$C + AE + BF$	DG
D	D	$AB + CG + EF$
E	$E + AC + BG$	DF
F	$F + BC + AG$	DE
G	$G + BE + AF$	CD

Por lo tanto, se ha aislado el efecto principal de D y todas sus interacciones de dos factores. En general, si a un diseño fraccionado de resolución III o mayor se le agrega una fracción adicional con los signos de *un solo factor* invertidos, entonces el diseño combinado producirá las estimaciones del efecto principal de ese factor y sus interacciones de dos factores.

Suponga ahora que a un diseño fraccionado de resolución III se le agrega una segunda fracción en la que los *signos de todos los factores están invertidos.* Este tipo de doblez (llamado en ocasiones **doblez completo** o reflexión) rompe los vínculos de alias entre los efectos principales y las interacciones de dos factores. Es decir, puede usarse el **diseño combinado** para estimar todos los efectos principales quitados de todas las interacciones de dos factores. En el siguiente ejemplo se ilustra la técnica.

EJEMPLO 8-7 ··

Un analista de desempeño humano conduce un experimento para estudiar el tiempo de enfoque del ojo y ha construido un aparato en el que pueden controlarse varios factores durante la prueba. Los factores que considera importantes inicialmente son la agudeza o claridad visual (A), la distancia del objetivo al ojo (B), la forma del objetivo (C), el nivel de iluminación (D), el tamaño del objetivo (E), la densidad del objetivo (F) y el sujeto (G). Se consideran dos niveles de cada factor. El analista sospecha que sólo algunos de estos siete factores son de importancia principal y que pueden omitirse las interacciones de órde-

Tabla 8-21 Diseño 2^{7-4}_{III} para el experimento del tiempo de enfoque del ojo

	Diseño básico								
Corrida	A	B	C	$D=AB$	$E=AC$	$F=BC$	$G=ABC$		Tiempo
1	−	−	−	+	+	+	−	*def*	85.5
2	+	−	−	−	−	+	+	*afg*	75.1
3	−	+	−	−	+	−	+	*beg*	93.2
4	+	+	−	+	−	−	−	*abd*	145.4
5	−	−	+	+	−	−	+	*cdg*	83.7
6	+	−	+	−	+	−	−	*ace*	77.6
7	−	+	+	−	−	+	−	*bcf*	95.0
8	+	+	+	+	+	+	+	*abcdefg*	141.8

nes superiores entre los factores. Con base en este supuesto, el analista decide correr un experimento de tamizado para identificar los factores más importantes para después enfocar el estudio en los mismos. Para explorar estos siete factores, el experimentador corre las combinaciones de tratamientos del diseño 2^{7-4}_{III} de la tabla 8-19 de manera aleatoria, obteniendo los tiempos de enfoque en milisegundos, como se muestra en la tabla 8-21.

A partir de estos datos pueden estimarse siete efectos principales y sus alias. Por la ecuación 8-1 se observa que los efectos y sus alias son

$$\begin{aligned}
\ell_A &= 20.63 \rightarrow A+BD+CE+FG\\
\ell_B &= 38.38 \rightarrow B+AD+CF+EG\\
\ell_C &= -0.28 \rightarrow C+AE+BF+DG\\
\ell_D &= 28.88 \rightarrow D+AB+CG+EF\\
\ell_E &= -0.28 \rightarrow E+AC+BG+DF\\
\ell_F &= -0.63 \rightarrow F+BC+AG+DE\\
\ell_G &= -2.43 \rightarrow G+CD+BE+AF
\end{aligned}$$

Por ejemplo,

$$\ell_A = \tfrac{1}{4}(-85.5+75.1-93.2+145.4-83.7+77.6-95.0+141.8) = 20.63$$

Los tres efectos más grandes son ℓ_A, ℓ_B y ℓ_D. La interpretación más simple de los datos es que los efectos principales de A, B y D son todos significativos. Sin embargo, esta interpretación no es única, ya que otra conclusión lógica sería que A, B y la interacción AB, o quizá B, D y la interacción BD, o tal vez A, D y la interacción AD son los verdaderos efectos.

Observe que ABD es una palabra en la relación de definición de este diseño. Por lo tanto, este diseño 2^{7-4}_{III} no se proyecta en un factorial 2^3 en ABD; en cambio, se proyecta en dos réplicas de un diseño 2^{3-1}, como se ilustra en la figura 8-23. Puesto que el diseño 2^{3-1} es de resolución III, A será alias de BD, B será alias de AD y D será alias de AB, por lo que no es posible separar las interacciones de los efectos principales. En este caso, quizá el analista haya tenido mala suerte. Si hubiera asignado el nivel de iluminación a C en lugar de a D, el diseño se habría proyectado en un diseño 2^3 completo, y la interpretación podría haber sido más sencilla.

Para separar los efectos principales y las interacciones de dos factores, se corre una segunda fracción con todos los signos invertidos. Este doblez del diseño se muestra en la tabla 8-22, junto con las respuestas observadas. Note que cuando se hace el doblez de un diseño de resolución III de esta manera, de hecho se

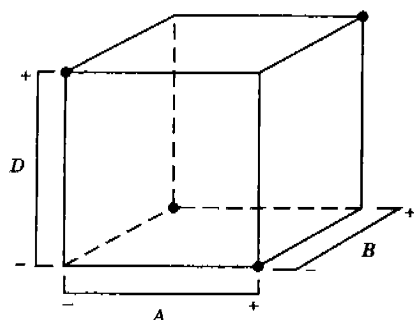


Figura 8-23 El diseño 2^{7-4}_{III} proyectado en dos réplicas de un diseño 2^{3-1}_{III} en A, B y D.

cambian los signos de los generadores que tienen un número impar de letras. Los efectos estimados por esta fracción son

$$\ell'_A = -17.68 \rightarrow A - BD - CE - FG$$
$$\ell'_B = 37.73 \rightarrow B - AD - CF - EG$$
$$\ell'_C = -3.33 \rightarrow C - AE - BF - DG$$
$$\ell'_D = 29.88 \rightarrow D - AB - CG - EF$$
$$\ell'_E = 0.53 \rightarrow E - AC - BG - DF$$
$$\ell'_F = 1.63 \rightarrow F - BC - AG - DE$$
$$\ell'_G = 2.68 \rightarrow G - CD - BE - AF$$

Al combinar esta segunda fracción con la original se obtienen las siguientes estimaciones de los efectos:

i	De $\frac{1}{2}(\ell_i + \ell'_i)$	De $\frac{1}{2}(\ell_i - \ell'_i)$
A	$A =$ 1.48	$BD + CE + FG =$ 19.15
B	$B =$ 38.05	$AD + CF + EG =$ 0.33
C	$C =$ −1.80	$AE + BF + DG =$ 1.53
D	$D =$ 29.38	$AB + CG + EF =$ −0.50
E	$E =$ 0.13	$AC + BG + DF =$ −0.40
F	$F =$ 0.50	$BC + AG + DE =$ −1.53
G	$G =$ 0.13	$CD + BE + AF =$ −2.55

Tabla 8-22 Un doblez del diseño 2^{7-4}_{III} en el experimento del tiempo de enfoque del ojo

	Diseño básico								
Corrida	A	B	C	$D = -AB$	$E = -AC$	$F = -BC$	$G = ABC$		Tiempo
1	+	+	+	−	−	−	+	*abcg*	91.3
2	−	+	+	+	+	−	−	*bcde*	136.7
3	+	−	+	+	−	+	−	*acdf*	82.4
4	−	−	+	−	+	+	+	*cefg*	73.4
5	+	+	−	−	+	+	−	*abef*	94.1
6	−	+	−	+	−	+	+	*bdfg*	143.8
7	+	−	−	+	+	−	+	*adeg*	87.3
8	−	−	−	−	−	−	−	(1)	71.9

Los dos efectos más grandes son B y D. Además, el tercer efecto más grande es $BD + CE + FG$, por lo que parece razonable atribuir esto a la interacción BD. El analista usó los dos factores, distancia (B) y nivel de iluminación (D), en experimentos subsecuentes con los demás factores A, C, E y F en ajustes estándar, y verificó los resultados obtenidos aquí. Decidió usar los sujetos como bloques en estos nuevos experimentos en lugar de ignorar el efecto potencial del sujeto debido a que fue necesario utilizar varios sujetos diferentes para completar el experimento.

La relación de definición para un diseño de doblez

La combinación de diseños factoriales fraccionados por medio de un doblez, como la que se hizo en el ejemplo 8-7, es una técnica muy útil. Con frecuencia es de interés conocer la relación de definición del diseño combinado. Puede determinarse con facilidad. Cada fracción separada tendrá $L + U$ palabras usadas como generadores: L palabras con el mismo signo y U palabras con signos diferentes. En el diseño combinado se usarán $L + U - 1$ palabras como generadores. Éstas serán las L palabras con el mismo signo y las $U - 1$ palabras que constan de productos pares independientes de las palabras que tienen signos diferentes. (Los **productos pares** son las palabras tomadas de dos en dos, de cuatro en cuatro, etcétera.)

Para ilustrar este procedimiento, considere el diseño del ejemplo 8-7. Para la primera fracción, los generadores son

$$I = ABD, \quad I = ACE, \quad I = BCF \quad \text{e} \quad I = ABCG$$

y para la segunda fracción son

$$I = -ABD, \quad I = -ACE, \quad I = -BCF \quad \text{e} \quad I = ABCG$$

Observe que en la segunda fracción se han intercambiado los signos de los generadores con un número impar de letras. Asimismo, observe que $L + U = 1 + 3 = 4$. El diseño combinado tendrá $I = ABCG$ (la palabra con el mismo signo) como generador y dos palabras que son productos pares independientes de las palabras con signos diferentes. Por ejemplo, tómese $I = ABD$ e $I = ACE$; entonces $I = (ABD)(ACE) = BCDE$ es un generador del diseño combinado. Asimismo, tómese $I = ABD$ e $I = BCF$; entonces $I = (ABD)(BCF) = ACDF$ es un generador del diseño combinado. La relación de definición completa para el diseño combinado es

$$I = ABCG = BCDE = ACDF = ADEG = BDFG = ABEF = CEFG$$

Diseños de Plackett-Burman

Estos diseños, atribuidos a Plackett y Burman, son diseños factoriales fraccionados de dos niveles para estudiar $k = N - 1$ variables en N corridas, donde N es un múltiplo de 4. Si N es una potencia de 2, estos diseños son idénticos a los que se presentaron anteriormente en esta sección. Sin embargo, para $N = 12, 20, 24, 28$ y 36, los diseños de Plackett-Burman en ocasiones son de interés. Puesto que estos diseños no pueden representarse como cubos, en ocasiones se les llama **diseños no geométricos.**

En la mitad superior de la tabla 8-23 se presentan los renglones de signos positivos y negativos que se usan para construir los diseños de Plackett-Burman para $N = 12, 20, 24$ y 36, mientras que en la mitad inferior de la tabla se presentan los bloques de signos positivos y negativos para construir el diseño para $N = 28$. Los diseños para $N = 12, 20, 24$ y 36 se obtienen escribiendo el renglón apropiado de la tabla 8-23 como una columna (o renglón). Entonces se genera una segunda columna (o renglón) a partir de la primera moviendo los elementos de la columna (o renglón) hacia abajo (o hacia la derecha) una posición y colocando el último elemento en la primera posición. Una tercera columna (o renglón) se produce a partir de la segunda de manera similar, y el proceso se continúa hasta que se genera la columna (o renglón) k.

Tabla 8-23 Signos positivos y negativos para los diseños de Plackett-Burman

$k = 11$, $N = 12$ ++−+++−−−+−		
$k = 19$, $N = 20$ ++−−++++−+−+−−−−++−		
$k = 23$, $N = 24$ +++++−+−++−−++−−+−+−−−−		
$k = 35$, $N = 36$ −+−+++−−−+++++−+++−−+−−−−+−+−++−−+−		
	$k = 27$, $N = 28$	
+−++++−−−	−+−−−+−−+	++−+−++−+
++−+++−−−	−−++−−+−−	−++++−++−
−++++ +−−−	+−−−+−−+−	+−+−++−++
−−−+−++++	−−+−+−−−+	+−+++−+−+
−−−++−+++	+−−−−++−−	++−−++++−
−−−−+++++	−+−+−−−+−	−+++−+−++
+++−−−+−+	−−+−−+−+−	+−++−+++−
+++−−−++−	+−−+−−−−+	++−++−−++
+++−−−−++	−+−−+−+−−	−++−+++−+

Después se agrega un renglón de signos negativos, completándose así el diseño. Para $N = 28$, los tres bloques X, Y y Z se apuntan en el orden

$$\begin{matrix} X & Y & Z \\ Z & X & Y \\ Y & Z & X \end{matrix}$$

y se agrega un renglón de signos negativos a estos 27 renglones. El diseño para $N = 12$ corridas y $k = 11$ factores se muestra en la tabla 8-24.

Los diseños no geométricos de Plackett-Burman para $N = 12, 20, 24, 28$ y 36 tienen estructuras de los alias muy intrincadas. Por ejemplo, en el diseño de 12 corridas, todos los efectos principales son **alias parciales** de cada una de las interacciones de dos factores en los que no están incluidos. Por ejemplo, la interacción AB es alias de los nueve efectos principales $C, D, ..., K$. Además, cada uno de los efectos principales son **alias parciales** de 45 interacciones de dos factores. En diseños más grandes, la situación es todavía más compleja. Se recomienda al experimentador usar estos diseños con *mucho cuidado.*

Tabla 8-24 Diseño de Plackett-Burman para $N = 12$, $k = 11$

Corrida	A	B	C	D	E	F	G	H	I	J	K
1	+	−	+	−	−	−	+	+	+	−	+
2	+	+	−	+	−	−	−	+	+	+	−
3	−	+	+	−	+	−	−	−	+	+	+
4	+	−	+	+	−	+	−	−	−	+	+
5	+	+	−	+	+	−	+	−	−	−	+
6	+	+	+	−	+	+	−	+	−	−	−
7	−	+	+	+	−	+	+	−	+	−	−
8	−	−	+	+	+	−	+	+	−	+	−
9	−	−	−	+	+	+	−	+	+	−	+
10	+	−	−	−	+	+	+	−	+	+	−
11	−	+	−	−	−	+	+	+	−	+	+
12	−	−	−	−	−	−	−	−	−	−	−

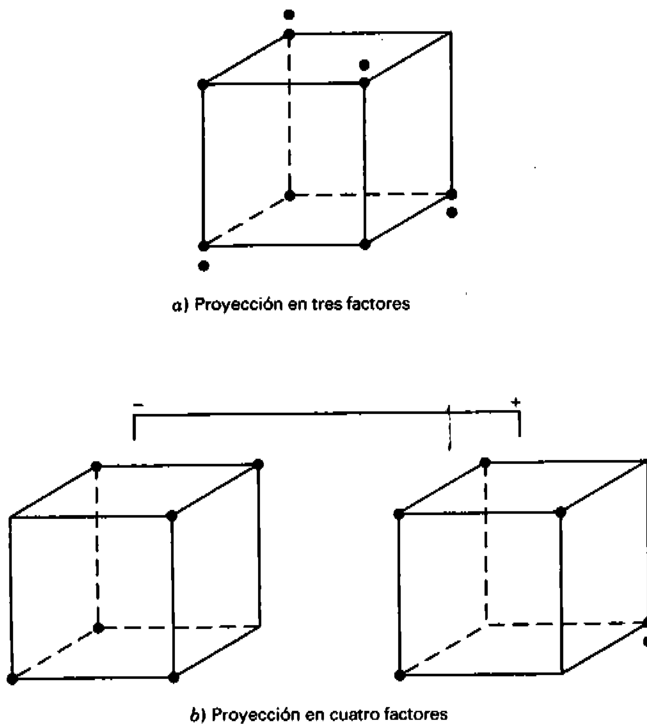


Figura 8-24 Proyección del diseño de Plackett-Burman de 12 corridas en diseños de tres y cuatro factores.

Las propiedades proyectivas de los diseños no geométricos de Plackett-Burman no son avasalladoramente atractivas. Por ejemplo, considere el diseño de 12 corridas de la tabla 8-24. Este diseño se proyectará en tres réplicas de un diseño 2^2 completo en dos cualesquiera de los 11 factores originales. Sin embargo, en tres factores, el diseño proyectado es un diseño 2^3 completo más un factorial fraccionado 2_{III}^{3-1} (ver la figura 8-24*a*). Por lo tanto, el diseño de Plackett-Burman de resolución III tiene **proyectividad** 3, lo cual significa que se plegará en un diseño factorial completo en cualquier subconjunto de tres factores. El diseño 2_{III}^{k-p} sólo tiene proyectividad 2. Las proyecciones de cuatro dimensiones se muestran en la figura 8-24*b*. Observe que estas proyecciones de tres y cuatro factores no son diseños balanceados.

EJEMPLO 8-8 ....

Se ilustrarán algunas de las dificultades potenciales asociadas con los diseños de Plackett-Burman utilizando el diseño de 11 variables con 12 corridas y un conjunto de datos simulados. Se supondrá que el proceso tiene tres efectos principales significativos (*A, B, D*) y dos interacciones significativas de dos factores (*AB* y *AD*). El modelo es

$$y = 200 + 8x_1 + 10x_2 + 12x_4 - 12x_1x_2 + 9x_1x_4 + \varepsilon$$

donde cada x_i es una variable codificada definida en el intervalo −1, +1 y ε es un término NID(0, 9) del error aleatorio. Por lo tanto, tres de los $k = 11$ factores son grandes, y hay dos interacciones grandes; la situación no está fuera de razón.

En la tabla 8-25 se presenta el diseño de Plackett-Burman con 12 corridas y las respuestas simuladas. Este diseño luce diferente al diseño de 12 corridas de la tabla 8-24 porque se construyó utilizando el ren-

Tabla 8-25 Diseño de Plackett-Burman para el ejemplo 8-8

Corrida	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*J*	*K*	*L*	Respuesta
1	+	+	−	+	+	+	−	−	−	+	−	231
2	−	+	+	−	+	+	+	−	−	−	+	207
3	+	−	+	+	−	+	+	+	−	−	−	230
4	−	+	−	+	+	−	+	+	+	−	−	217
5	−	−	+	−	+	+	−	+	+	+	−	175
6	−	−	−	+	−	+	+	−	+	+	+	176
7	+	−	−	−	+	−	+	+	−	+	+	183
8	+	+	−	−	−	+	−	+	+	−	+	185
9	+	+	+	−	−	−	+	−	+	+	−	181
10	−	+	+	+	−	−	−	+	−	+	+	220
11	+	−	+	+	+	−	−	−	+	−	+	229
12	−	−	−	−	−	−	−	−	−	−	−	168

glón de signos para $k = 11, N = 12$ de la tabla 8-23 como renglón. En la tabla 8-26 se muestran las estimaciones de los efectos. Observe que hay siete efectos grandes: *A, B, C, D, E, J* y *K* (y, desde luego, sus alias). No es evidente de inmediato que algunos de estos efectos podrían ser interacciones. Parte de esta ambigüedad podría resolverse haciendo el doblez del diseño. Con esto por lo general se resolverán los efectos principales, pero con frecuencia sigue dejando al experimentador con la incertidumbre acerca de los efectos de las interacciones.

....

La dificultad para interpretar un diseño de Plackett-Burman, ilustrada en el ejemplo anterior, ocurre con mucha frecuencia en la práctica. Si la elección está entre un diseño geométrico 2_{III}^{11-7} con 16 corridas o un diseño de Plackett-Burman con 12 corridas que quizá tenga que doblarse (para lo cual se requerirían 24 corridas), el diseño geométrico puede resultar una mejor elección. Para mayores detalles ver Montgomery, Borror y Stanley [81]. Bajo ciertas condiciones, los alias de un diseño no geométrico de Plac-

Tabla 8-26 Estimaciones de los efectos, coeficientes de regresión y sumas de cuadrados del ejemplo 8-8

Variable[a]	Coeficiente de regresión	Efecto estimado	Suma de cuadrados
Promedio global	200.167		
A	6.333	12.667	481.333
B	6.667	13.333	533.333
C	6.833	12.667	560.333
D	17.000	34.000	3468.000
E	6.833	13.667	560.333
F	0.500	1.000	3.000
G	−1.167	−2.333	16.333
H	1.500	3.000	27.000
J	−6.333	−12.667	481.333
K	−5.833	−11.667	408.333
L	−0.167	−0.333	0.333

[a]Todos los efectos principales son alias parciales de 45 interacciones de dos factores.

kett-Burman pueden desenredarse utilizando técnicas de construcción de modelos de regresión. Esto se analiza en Hamada y Wu [53].

8-6 DISEÑOS DE RESOLUCIÓN IV Y V

Un diseño factorial fraccionado 2^{k-p} es de resolución IV si los efectos principales están separados de las interacciones de dos factores y algunas interacciones de dos factores son alias entre sí. Por lo tanto, si se suprimen las interacciones de tres factores y de órdenes superiores, los efectos principales pueden estimarse directamente en un diseño 2_{IV}^{k-p}. Un ejemplo es el diseño 2_{IV}^{6-2} de la tabla 8-10. Además, las dos fracciones combinadas del diseño 2_{III}^{7-4} del ejemplo 8-7 producen un diseño 2_{IV}^{7-3}.

Cualquier diseño 2_{IV}^{k-p} debe incluir al menos $2k$ corridas. A los diseños de resolución IV que contienen exactamente $2k$ corridas se les llama **diseños mínimos.** Los diseños de resolución IV pueden obtenerse a partir de diseños de resolución III por el proceso de **doblado.** Recuerde que para hacer el doblez de un diseño 2_{III}^{k-p}, simplemente se agrega a la fracción original una segunda fracción con todos los signos invertidos. Entonces los signos positivos en la columna identidad I de la primera fracción podrían intercambiarse en la segunda fracción, y el factor $(k + 1)$-ésimo podría asociarse con esta columna. El resultado es un diseño factorial fraccionado 2_{IV}^{k+1-p}. El proceso se muestra en la tabla 8-27 para el diseño 2_{III}^{3-1}. Es sencillo verificar que el diseño resultante es un diseño 2_{IV}^{4-1} con la relación de definición $I = ABCD$.

También es posible hacer el doblez de diseños de resolución IV para separar las interacciones de dos factores que son alias entre sí. Montgomery y Runger [83c] hacen notar que un experimentador puede tener varios objetivos al hacer el doblez de un diseño de resolución IV, como 1) romper tantas cadenas de alias de interacciones de dos factores como sea posible, 2) romper las interacciones de dos factores en una cadena de alias específica, o 3) romper las interacciones de dos factores que incluyen un factor específico. Una manera de hacer el doblez de un diseño de resolución IV es corriendo una segunda fracción en la que se invierte el signo de todos los generadores del diseño que tienen un número impar de letras. Para ilustrar, considere el diseño 2_{IV}^{6-2} usado en el experimento del moldeo por inyección del ejemplo 8-4. Los generadores del diseño de la tabla 8-10 son $I = ABCE$ e $I = BCDF$. La segunda fracción usaría los

Tabla 8-27 Diseño 2_{IV}^{4-1} obtenido por doblez

D *I*	*A*	*B*	*C*
Diseño 2_{III}^{3-1} orginal con $I = ABC$			
+	−	−	+
+	+	−	−
+	−	+	−
+	+	+	+
Segundo diseño 2_{III}^{3-1} con los signos intercambiados			
−	+	+	−
−	−	+	+
−	+	−	+
−	−	−	−

generadores $I = -ABCE$ e $I = -BCDF$, y el generador único para el diseño combinado sería $I = ADEF$. Por lo tanto, el diseño combinado sigue siendo un diseño factorial fraccionado de resolución IV. Sin embargo, las relaciones de los alias serán mucho más sencillas que en el diseño 2_{IV}^{6-2} original. De hecho, las únicas interacciones de dos factores que tendrán alias son $AD = EF$, $AE = DF$ y $AF = DE$. Todas las demás interacciones de dos factores pueden estimarse a partir del diseño combinado.

Como otro ejemplo, considere el diseño 2_{IV}^{8-3} con 32 corridas. La tabla 8-14 indica que el mejor conjunto de generadores para este diseño es $I = ABCF$, $I = ABDG$ e $I = BCDEH$. En la tabla XII(m) del apéndice se muestran los alias para este diseño. Observe que hay seis pares de interacciones de dos factores y un grupo de tres interacciones de dos factores que son alias. Si se hace el doblez de este diseño, la segunda fracción tendría los generadores $I = -ABCF$, $I = -ABDG$ e $I = BCDEH$. El diseño combinado tiene los generadores $I = CDFG$ e $I = BCDEH$, y la relación de definición completa es

$$I = CDFG = BCDEH = BEFGH$$

El diseño combinado es de resolución IV, pero las únicas interacciones de dos factores que siguen teniendo alias son $CD = FG$, $CF = DG$ y $CG = DF$. Se trata de una simplificación considerable de los alias de la fracción original.

Observe que cuando se empieza con un diseño de resolución III, el procedimiento de doblez garantiza que el diseño combinado será de resolución IV, con lo cual se asegura que todos los efectos principales pueden separarse de sus alias en interacciones de dos factores. Cuando se hace el doblez de un diseño de resolución IV, no necesariamente se separarán *todas* las interacciones de dos factores. De hecho, si la fracción *original* tiene una estructura de los alias con más de dos interacciones de dos factores en cualquier cadena de alias, el doblez no separará completamente todas las interacciones de dos factores. Ambos ejemplos anteriores, el 2_{IV}^{6-2} y el 2_{IV}^{8-3}, tienen al menos una de tales cadenas de alias de interacciones de dos factores. Montgomery y Runger [83c] dan una tabla de diseños hechos doblez recomendados para fracciones de resolución IV con $6 \leq k \leq 10$ factores.

Los diseños de resolución V son factoriales fraccionados en los que los efectos principales y las interacciones de dos factores no tienen como alias otros efectos principales u otras interacciones de dos factores. Estos diseños son muy poderosos, permitiendo la estimación única de todos los efectos principales y las interacciones de dos factores, siempre que todas las interacciones de tres factores y de órdenes superiores sean insignificantes. La palabra más pequeña de la relación de definición de tal diseño debe tener cinco letras. El diseño 2^{5-1} con la relación de definición $I = ABCDE$ es de resolución V. Otro ejemplo es el diseño 2_{V}^{8-2} con las relaciones de definición $I = ABCDG$ e $I = ABEFH$. Ejemplos adicionales de estos diseños se presentan en Box y Hunter [17c].

Debido a que los diseños estándar de resolución V son diseños grandes cuando el número de factores es moderadamente grande, existe cierto interés práctico en los diseños factoriales fraccionados irregulares de resolución V. Se cuenta con diseños útiles para $4 \leq k \leq 9$ factores. El diseño de 24 corridas para $k = 5$ factores se muestra en la tabla 8-28. Puesto que se trata de un diseño de resolución V, es posible estimar los cinco efectos principales y las 10 interacciones de dos factores, suponiendo que las interacciones de tres factores y órdenes superiores son insignificantes. El diseño para $k = 4$ factores tiene 12 corridas y se comenta en el problema 8-22. Para $k = 6$, 7 y 8, estos diseños tienen 48 corridas, y el diseño de nueve factores tiene 96 corridas. El paquete de software *Design-Expert* contiene todos estos diseños.

Por último, cabe señalar que un doblez completo de un diseño de resolución IV o V suele ser innecesario. En general, sólo hay una o dos (o muy pocas) interacciones con alias que son de interés potencial. Los alias de estas interacciones pueden por lo general separarse agregando un número pequeño de corri-

Tabla 8-28 La fracción irregular de resolución V para cinco factores en 24 corridas

A	*B*	*C*	*D*	*E*
−	−	−	−	−
−	+	−	−	−
+	+	−	−	−
−	−	+	−	−
+	−	+	−	−
−	+	+	−	−
−	−	−	+	−
+	−	−	+	−
+	+	−	+	−
+	−	+	+	−
−	+	+	+	−
+	+	+	+	−
−	−	−	−	+
+	−	−	−	+
+	+	−	−	+
+	−	+	−	+
−	+	+	−	+
+	+	+	−	+
−	−	−	+	+
−	+	−	+	+
+	+	−	+	+
−	−	+	+	+
+	−	+	+	+
−	+	+	+	+

das a la fracción original. Esta técnica se denomina en ocasiones **doblez parcial.** Para formarse una idea de cómo se hace esto, referirse al ejemplo 10-5 y al material suplementario del texto de este capítulo.

8-7 RESUMEN

En este capítulo se introdujo el diseño factorial fraccionado 2^{k-p}. Se ha hecho hincapié en el uso de estos diseños en experimentos de tamizado para identificar de manera rápida y eficaz el subconjunto de factores que están activos, así como para proporcionar cierta información sobre las interacciones. La propiedad de proyección de estos diseños hace posible en muchos casos examinar los factores activos con mayor detalle. El ensamblaje secuencial de estos diseños por medio de un doblez es una manera muy eficaz de obtener información adicional acerca de las interacciones que pueden identificarse como de posible importancia en un experimento inicial.

En la práctica, los diseños factoriales fraccionados 2^{k-p} con $N = 4, 8, 16$ y 32 corridas son muy útiles. En la tabla 8-29 se resumen estos diseños, identificando cuántos factores pueden usarse con cada diseño para obtener diferentes tipos de experimentos de tamizado. Por ejemplo, el diseño de 16 corridas es un factorial completo para 4 factores, una fracción un medio para 5 factores, una fracción de resolución IV para 6 u 8 factores y una fracción de resolución III para 9 a 15 factores. Todos estos diseños pueden construirse utilizando los métodos explicados en este capítulo, y muchas de sus estructuras de los alias se muestran en la tabla XII del apéndice.

Tabla 8-29 Diseños factoriales y factoriales fraccionados útiles del sistema 2^{k-p}. Los números en las celdas son el número de factores del experimento

	Número de corridas			
Tipo de diseño	4	8	16	32
Factorial completo	2	3	4	5
Fracción un medio	3	4	5	6
Fracción de resolución IV	—	4	6–8	7–16
Fracción de resolución III	3	5–7	9–15	17–31

8-8 PROBLEMAS

8-1. Suponga que en el experimento del desarrollo del proceso químico descrito en el problema 6-7 sólo pudo correrse una fracción un medio del diseño 2^4. Construir el diseño y llevar a cabo el análisis estadístico utilizando los datos de la réplica I.

8-2. Suponga que en el problema 6-15 sólo pudo correrse una fracción un medio del diseño 2^4. Construir el diseño y llevar a cabo el análisis utilizando los datos de la réplica I.

8-3. Considere el experimento del grabado con plasma del problema 6-18. Suponga que sólo pudo correrse una fracción un medio del diseño. Establecer el diseño y analizar los datos.

8-4. En el problema 6-21 se describe el estudio para mejorar un proceso durante la manufactura de un circuito integrado. Suponga que sólo pudieron hacerse ocho corridas de este proceso. Establecer un diseño 2^{5-2} apropiado y encontrar la estructura de los alias. Utilizar las observaciones apropiadas del problema 6-21 como las observaciones de este diseño y estimar los efectos de los factores. ¿Qué conclusiones pueden sacarse?

8-5. ***Continuación del problema 8-4.*** Suponga que ha hecho las ocho corridas del diseño 2^{5-2} del problema 8-4. ¿Qué corridas adicionales se necesitarían para identificar los efectos de los factores que son de interés? ¿Cuáles son las relaciones de los alias en el diseño combinado?

8-6. R.D. Snee ("Experimentación con un número grande de variables", en *Experiments in Industry: Design, Analysis and Interpretation of Results*, de R.D. Snee, L.B. Hare y J.B. Trout, editores, ASQC) describe un experimento en el que se usó un diseño 2^{5-1} con $I = ABCDE$ para investigar los efectos de cinco factores sobre el color de un producto químico. Los factores son A = solvente/reactivo, B = catalizador/reactivo, C = temperatura, D = pureza del reactivo y E = pH del reactivo. Los resultados obtenidos fueron los siguientes:

$e = -0.63$	$d = 6.79$
$a = 2.51$	$ade = 5.47$
$b = -2.68$	$bde = 3.45$
$abe = 1.66$	$abd = 5.68$
$c = 2.06$	$cde = 5.22$
$ace = 1.22$	$acd = 4.38$
$bce = -2.09$	$bcd = 4.30$
$abc = 1.93$	$abcde = 4.05$

a) Construir una gráfica de probabilidad normal de los efectos. ¿Qué efectos parecen estar activos?

b) Calcular los residuales. Construir una gráfica de probabilidad normal de los residuales y graficar los residuales contra los valores ajustados. Comentar las gráficas.

c) Si algunos de los factores son insignificantes, plegar el diseño 2^{5-1} a un diseño factorial completo en los factores activos. Comentar el diseño resultante e interpretar los resultados.

8-7. En un artículo de J.J. Pignatiello, Jr. y J.S. Ramberg del *Journal of Quality Technology* (vol. 17, pp. 198-206) se describe el uso de un diseño factorial fraccionado con réplicas para investigar el efecto de cinco factores sobre la altura libre de los resortes de hojas utilizados en una aplicación automotriz. Los factores son A = tem-

peratura del horno, B = tiempo de calentamiento, C = tiempo de transferencia, D = tiempo de retención y E = temperatura del aceite de templado. Los datos se presentan a continuación:

A	B	C	D	E	Altura libre		
−	−	−	−	−	7.78	7.78	7.81
+	−	−	+	−	8.15	8.18	7.88
−	+	−	+	−	7.50	7.56	7.50
+	+	−	−	−	7.59	7.56	7.75
−	−	+	+	−	7.54	8.00	7.88
+	−	+	−	−	7.69	8.09	8.06
−	+	+	−	−	7.56	7.52	7.44
+	+	+	+	−	7.56	7.81	7.69
−	−	−	−	+	7.50	7.25	7.12
+	−	−	+	+	7.88	7.88	7.44
−	+	−	+	+	7.50	7.56	7.50
+	+	−	−	+	7.63	7.75	7.56
−	−	+	+	+	7.32	7.44	7.44
+	−	+	−	+	7.56	7.69	7.62
−	+	+	−	+	7.18	7.18	7.25
+	+	+	+	+	7.81	7.50	7.59

a) Escribir la estructura de los alias de este diseño. ¿Qué resolución tiene este diseño?
b) Analizar los datos. ¿Qué factores influyen en la altura libre promedio?
c) Calcular el rango y la desviación estándar de la altura libre para cada corrida. ¿Hay algún indicio de que cualquiera de estos factores afecta la variabilidad de la altura libre?
d) Analizar los residuales de este experimento y comentar los resultados.
e) ¿Este diseño es el mejor posible para cinco factores en 16 corridas? Específicamente, ¿es posible encontrar un diseño fraccionado para cinco factores en 16 corridas con una resolución más alta que la de este diseño?

8-8. En un artículo de *Industrial and Engineering Chemistry* ("Información adicional acerca de la planeación de experimentos para aumentar la eficiencia de la investigación") se utiliza un diseño 2^{5-2} para investigar el efecto de A = temperatura de condensación, B = cantidad del material 1, C = volumen del solvente, D = tiempo de condensación y E = cantidad del material 2 sobre el rendimiento. Los resultados obtenidos son los siguientes:

$$e = 23.2 \quad ad = 16.9 \quad cd = 23.8 \quad bde = 16.8$$
$$ab = 15.5 \quad bc = 16.2 \quad ace = 23.4 \quad abcde = 18.1$$

a) Verificar que los generadores que se utilizaron en el diseño fueron $I = ACE$ e $I = BDE$.
b) Apuntar la relación de definición completa y los alias de este diseño.
c) Estimar los efectos principales.
d) Elaborar la tabla del análisis de varianza. Verificar que las interacciones AB y AD están disponibles para usarlas como error.
e) Graficar los residuales contra los valores ajustados. Construir también la gráfica de probabilidad normal de los residuales. Comentar los resultados.

8-9. Considere el experimento con el resorte de hojas del problema 8-7. Suponga que el factor E (temperatura del aceite de templado) es muy difícil de controlar durante la manufactura. ¿Cuál sería el ajuste de los factores A, B, C y D para reducir la variabilidad de la altura libre tanto como sea posible, independientemente de la temperatura del aceite de templado usada?

8-10. Construir un diseño 2^{7-2} seleccionando dos interacciones de dos factores como los generadores independientes. Apuntar la estructura de los alias completa de este diseño. Delinear la tabla del análisis de varianza. ¿Cuál es la resolución de este diseño?

8-11. Considere el diseño 2^5 del problema 6-21. Suponga que sólo pudo correrse una fracción un medio. Además se requirieron dos días para hacer las 16 observaciones, y fue necesario confundir el diseño 2^{5-1} en dos bloques. Construir el diseño y analizar los datos.

8-12. Analizar los datos del problema 6-23 como si provinieran de un diseño 2_{IV}^{4-1} con $I = ABCD$. Proyectar el diseño en un factorial completo en el subconjunto de los cuatro factores originales que parecen ser significativos.

8-13. Repetir el problema 8-12 utilizando $I = -ABCD$. ¿El uso de la fracción alterna modifica la interpretación de los datos?

8-14. Proyectar el diseño 2_{IV}^{4-1} del ejemplo 8-1 en dos réplicas de un diseño 2^2 en los factores A y B. Analizar los datos y sacar conclusiones.

8-15. Construir un diseño 2_{III}^{6-3}. Determinar los efectos que pueden estimarse si se corre una segunda fracción de este diseño con todos los signos invertidos.

8-16. Considere el diseño 2_{III}^{6-3} del problema 8-15. Determinar los efectos que pueden estimarse si se corre una segunda fracción de este diseño con los signos del factor A invertidos.

8-17. Hacer el doblez del diseño 2_{III}^{7-4} de la tabla 8-19 para producir un diseño de ocho factores. Verificar que el diseño resultante sea 2_{IV}^{8-4}. ¿Se trata de un diseño mínimo?

8-18. Hacer el doblez de un diseño 2_{III}^{5-2} para producir un diseño de seis factores. Verificar que el diseño resultante sea 2_{IV}^{6-2}. Comparar este diseño con el diseño 2_{IV}^{6-2} de la tabla 8-10.

8-19. Un ingeniero industrial realiza un experimento utilizando un modelo de simulación Montecarlo de un sistema de inventario. Las variables independientes de su modelo son la cantidad del pedido (A), el punto de un nuevo pedido (B), el costo de organización (C), el costo del refrendo de pedidos (D) y la tarifa de transportación (E). La variable de respuesta es el costo anual promedio. Para ahorrar tiempo de computadora, el ingeniero decide investigar estos factores utilizando un diseño 2_{III}^{5-2} con $I = ABD$ e $I = BCE$. Los resultados que obtiene son $de = 95$, $ae = 134$, $b = 158$, $abd = 190$, $cd = 92$, $ac = 187$, $bce = 155$ y $abcde = 185$.

   a) Verificar que las combinaciones de tratamientos dadas sean correctas. Estimar los efectos suponiendo que las interacciones de tres factores y de órdenes superiores son insignificantes.

   b) Suponga que se agrega una segunda fracción a la primera, por ejemplo, $ade = 136$, $e = 93$, $ab = 187$, $bd = 153$, $acd = 139$, $c = 99$, $abce = 191$ y $bcde = 150$. ¿Cómo se obtuvo esta segunda fracción? Incorporar estos datos a la fracción original y estimar los efectos.

   c) Suponga que se corrió la fracción $abc = 189$, $ce = 96$, $bcd = 154$, $acde = 135$, $abe = 193$, $bde = 152$, $ad = 137$ y $(1) = 98$. ¿Cómo se obtuvo esta fracción? Incorporar estos datos en la fracción original y estimar los efectos.

8-20. Construir un diseño 2^{5-1}. Indicar cómo puede correrse el diseño en dos bloques de ocho observaciones cada uno. ¿Alguno de los efectos principales o de las interacciones de dos factores están confundidos con los bloques?

8-21. Construir un diseño 2^{7-2}. Indicar cómo puede correrse el diseño en cuatro bloques de ocho observaciones cada uno. ¿Alguno de los efectos principales o de las interacciones de dos factores están confundidos con los bloques?

8-22. ***Fracciones irregulares del diseño 2^k (John [61d]).*** Considere un diseño 2^4. Tienen que estimarse los cuatro efectos principales y las seis interacciones de dos factores, pero no puede correrse el factorial 2^4 completo. El tamaño del bloque más grande posible contiene 12 corridas. Estas 12 corridas pueden obtenerse de las cuatro réplicas un cuarto definidas por $I = \pm AB = \pm ACD = \pm BCD$ omitiendo la fracción principal. Indicar cómo pueden combinarse las tres fracciones 2^{4-2} restantes para estimar los efectos requeridos, suponiendo que las interacciones de tres factores y de órdenes superiores son insignificantes. Este diseño podría considerarse como una fracción tres cuartos.

8-23. Los ánodos de carbono utilizados en un proceso de fundición se fabrican en un horno anular. Se corre un experimento en el horno para determinar cuáles son los factores que influyen en el peso del material de empaque que se adhiere a los ánodos después de la cocción. Seis variables son de interés, cada una con dos niveles: A = relación paso/finos (0.45, 0.55), B = tipo de material de empaque (1, 2), C = temperatura del material de empaque (ambiente, 325°C), D = localización de la chimenea (adentro, afuera), E = temperatura del foso (ambiente, 195°C) y F = tiempo de retraso antes del empaque (cero, 24 horas). Se corre un diseño 2^{6-3} y se

obtienen tres réplicas en cada uno de los puntos del diseño. El peso del material de empaque adherido a los ánodos se mide en gramos. Los datos en el orden de las corridas son los siguientes: *abd* = (984, 826, 936); *abcdef* = (1275, 976, 1457); *be* = (1217, 1201, 890); *af* = (1474, 1164, 1541); *def* = (1320, 1156, 913); *cd* = (765, 705, 821); *ace* = (1338, 1254, 1294) y *bcf* = (1325, 1299, 1253). Se desea minimizar la cantidad de material de empaque adherido.

a) Verificar que las ocho corridas correspondan a un diseño 2^{6-3}_{III}. ¿Cuál es la estructura de los alias?
b) Usar el peso promedio como respuesta. ¿Qué factores parecen tener influencia?
c) Usar el rango de los pesos como respuesta. ¿Qué factores parecen tener influencia?
d) ¿Qué recomendaciones podrían hacerse a los ingenieros del proceso?

8-24. Se corrió un experimento de 16 corridas en una planta de manufactura de semiconductores para estudiar los efectos de seis factores sobre la curvatura o combadura de los dispositivos del sustrato producidos. Las seis variables y sus niveles se presentan a continuación:

Corrida	Temperatura de laminación (°C)	Tiempo de laminación (s)	Presión de laminación (tn)	Temperatura de cocción (°C)	Duración del ciclo de cocción (h)	Punto de rocío de la cocción (°C)
1	55	10	5	1580	17.5	20
2	75	10	5	1580	29	26
3	55	25	5	1580	29	20
4	75	25	5	1580	17.5	26
5	55	10	10	1580	29	26
6	75	10	10	1580	17.5	20
7	55	25	10	1580	17.5	26
8	75	25	10	1580	29	20
9	55	10	5	1620	17.5	26
10	75	10	5	1620	29	20
11	55	25	5	1620	29	26
12	75	25	5	1620	17.5	20
13	55	10	10	1620	29	20
14	75	10	10	1620	17.5	26
15	55	25	10	1620	17.5	20
16	75	25	10	1620	29	26

Se hicieron cuatro réplicas de cada corrida, y se hizo una medición de la combadura del sustrato. Los datos se presentan enseguida:

	Combadura por réplica (pulg/pulg)				Total	Media	Desviación
Corrida	1	2	3	4	(10^{-4} pulg/pulg)	(10^{-4} pulg/pulg)	estándar
1	0.0167	0.0128	0.0149	0.0185	629	157.25	24.418
2	0.0062	0.0066	0.0044	0.0020	192	48.00	20.976
3	0.0041	0.0043	0.0042	0.0050	176	44.00	4.083
4	0.0073	0.0081	0.0039	0.0030	223	55.75	25.025
5	0.0047	0.0047	0.0040	0.0089	223	55.75	22.410
6	0.0219	0.0258	0.0147	0.0296	920	230.00	63.639
7	0.0121	0.0090	0.0092	0.0086	389	97.25	16.029
8	0.0255	0.0250	0.0226	0.0169	900	225.00	39.42
9	0.0032	0.0023	0.0077	0.0069	201	50.25	26.725
10	0.0078	0.0158	0.0060	0.0045	341	85.25	50.341

	Combadura por réplica (pulg/pulg)				Total	Media	Desviación
Corrida	1	2	3	4	(10^{-4} pulg/pulg)	(10^{-4} pulg/pulg)	estándar
11	0.0043	0.0027	0.0028	0.0028	126	31.50	7.681
12	0.0186	0.0137	0.0158	0.0159	640	160.00	20.083
13	0.0110	0.0086	0.0101	0.0158	455	113.75	31.12
14	0.0065	0.0109	0.0126	0.0071	371	92.75	29.51
15	0.0155	0.0158	0.0145	0.0145	603	150.75	6.75
16	0.0093	0.0124	0.0110	0.0133	460	115.00	17.45

a) ¿Qué tipo de diseño utilizaron los experimentadores?
b) ¿Cuáles son las relaciones de los alias en este diseño?
c) ¿Alguna de las variables del proceso afecta la combadura promedio?
d) ¿Alguna de las variables del proceso afecta la variabilidad de las mediciones de la combadura?

Tabla 8-30 Datos para el problema 8-25

	A	*B*	*C*	*D*	*E*	*F*	Espesor del recubrimiento protector				
Corrida	Volumen	Lote	Tiempo, s	Velocidad	Aceleración	Cubierta	Izq.	Centro	Der.	Prom.	Rango
1	5	Lote 2	14	7350	5	Sin	4531	4531	4515	4525.7	16
2	5	Lote 1	6	7350	5	Sin	4446	4464	4428	4446	36
3	3	Lote 1	6	6650	5	Sin	4452	4490	4452	4464.7	38
4	3	Lote 2	14	7350	20	Sin	4316	4328	4308	4317.3	20
5	3	Lote 1	14	7350	5	Sin	4307	4295	4289	4297	18
6	5	Lote 1	6	6650	20	Sin	4470	4492	4495	4485.7	25
7	3	Lote 1	6	7350	5	Con	4496	4502	4482	4493.3	20
8	5	Lote 2	14	6650	20	Sin	4542	4547	4538	4542.3	9
9	5	Lote 1	14	6650	5	Sin	4621	4643	4613	4625.7	30
10	3	Lote 1	14	6650	5	Con	4653	4670	4645	4656	25
11	3	Lote 2	14	6650	20	Con	4480	4486	4470	4478.7	16
12	3	Lote 1	6	7350	20	Sin	4221	4233	4217	4223.7	16
13	5	Lote 1	6	6650	5	Con	4620	4641	4619	4626.7	22
14	3	Lote 1	6	6650	20	Con	4455	4480	4466	4467	25
15	5	Lote 2	14	7350	20	Con	4255	4288	4243	4262	45
16	5	Lote 2	6	7350	5	Con	4490	4534	4523	4515.7	44
17	3	Lote 2	14	7350	5	Con	4514	4551	4540	4535	37
18	3	Lote 1	14	6650	20	Sin	4494	4503	4496	4497.7	9
19	5	Lote 2	6	7350	20	Sin	4293	4306	4302	4300.3	13
20	3	Lote 2	6	7350	5	Sin	4534	4545	4512	4530.3	33
21	5	Lote 1	14	6650	20	Con	4460	4457	4436	4451	24
22	3	Lote 2	6	6650	5	Con	4650	4688	4656	4664.7	38
23	5	Lote 1	14	7350	20	Sin	4231	4244	4230	4235	14
24	3	Lote 2	6	7350	20	Con	4225	4228	4208	4220.3	20
25	5	Lote 1	14	7350	5	Con	4381	4391	4376	4382.7	15
26	3	Lote 2	6	6650	20	Sin	4533	4521	4511	4521.7	22
27	3	Lote 1	14	7350	20	Con	4194	4230	4172	4198.7	58
28	5	Lote 2	6	6650	5	Sin	4666	4695	4672	4677.7	29
29	5	Lote 1	6	7350	20	Con	4180	4213	4197	4196.7	33
30	5	Lote 2	6	6650	20	Con	4465	4496	4463	4474.7	33
31	5	Lote 2	14	6650	5	Con	4653	4685	4665	4667.7	32
32	3	Lote 2	14	6650	5	Sin	4683	4712	4677	4690.7	35

e) Si es importante reducir la combadura tanto como sea posible, ¿qué recomendaciones se harían?

8-25. Se usa un revestimiento por centrifugado para aplicar un recubrimiento fotoprotector en una oblea de silicio natural. Esta operación suele hacerse en las fases iniciales del proceso de fabricación de semiconductores, y el espesor promedio del recubrimiento protector y la variabilidad del espesor del mismo tienen un impacto importante en los pasos subsecuentes de manufactura. Seis variables se usan en el experimento. Las variables y sus niveles alto y bajo se presentan a continuación:

Factor	Nivel bajo	Nivel alto
Velocidad de centrifugado final	7350 rpm	6650 rpm
Índice de aceleración	5	20
Volumen de recubrimiento protector aplicado	3 cc	5 cc
Tiempo del centrifugado	14 s	6 s
Variación del lote del recubrimiento protector	Lote 1	Lote 2
Presión de descarga	Sin cubierta	Con cubierta

El experimentador decide usar un diseño 2^{6-1} y hacer tres lecturas del espesor del recubrimiento protector en cada oblea de prueba. Los datos se muestran en la tabla 8-30.

a) Verificar que se trata de un diseño 2^{6-1}. Discutir las relaciones de los alias de este diseño.
b) ¿Qué factores parecen afectar el espesor promedio del recubrimiento protector?
c) Considerando que el volumen del recubrimiento protector aplicado tiene un efecto reducido sobre el espesor promedio, ¿tiene esto alguna implicación práctica importante para los ingenieros del proceso?
d) Proyectar este diseño en un diseño menor que incluya únicamente los factores significativos. Presentar los resultados gráficamente. ¿Ayuda esto en la interpretación?
e) Usar el rango del espesor del recubrimiento protector como variable de respuesta. ¿Hay algún indicio de que alguno de estos factores afecte la variabilidad del espesor del recubrimiento protector?
f) ¿Dónde se recomendaría que corrieran el proceso los ingenieros?

8-26. Harry y Judy Peterson-Nedry (dos amigos del autor) son propietarios de un viñedo y una fábrica vinícola en Newberg, Oregon. Cultivan varias variedades de uvas y fabrican vino. Harry y Judy han usado diseños factoriales para el desarrollo de procesos y productos en el segmento de fabricación vinícola de su negocio. Este problema describe el experimento realizado para su Pinot Noir 1985. Originalmente se estudiaron ocho variables, las cuales se muestran en este experimento:

Variable	Nivel bajo (−)	Nivel alto (+)
A = Clon de Pinot Noir	Pommard	Wadenswil
B = Tipo de roble	Allier	Tronçais
C = Edad de la barrica	Vieja	Nueva
D = Levadura/contacto con la piel	Champagne	Montrachet
E = Vapores	Ninguno	Todos
F = Tostado de las barricas	Ligero	Medio
G = Racimos completos	Ninguno	10%
H = Temperatura de fermentación	Baja (75°F máx.)	Alta (92°F máx.)

Harry y Judy decidieron usar un diseño 2_{IV}^{8-4} con 16 corridas. El vino fue catado por un panel de expertos el 8 de marzo de 1986. Cada experto calificó las 16 muestras de vino catadas, siendo la calificación 1 la mejor. El diseño y los resultados del panel de catadores se muestra en la tabla 8-31.

a) ¿Cuáles son las relaciones de los alias en el diseño seleccionado por Harry y Judy?
b) Usar las calificaciones promedio ($\bar{y}$) como variable de respuesta. Analizar los datos y sacar conclusiones. Se encontrará útil examinar una gráfica de probabilidad normal de las estimaciones de los efectos.

Tabla 8-31 Diseño y resultados del experimento de la prueba del vino

	Variable								Calificaciones del panel					Resumen	
Corrida	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	HPN	JPN	CAL	DCM	RGB	$\bar{y}$	*s*
1	−	−	−	−	−	−	−	−	12	6	13	10	7	9.6	3.05
2	+	−	−	−	−	+	+	+	10	7	14	14	9	10.8	3.11
3	−	+	−	−	+	−	+	+	14	13	10	11	15	12.6	2.07
4	+	+	−	−	+	+	−	−	9	9	7	9	12	9.2	1.79
5	−	−	+	−	+	+	+	−	8	8	11	8	10	9.0	1.41
6	+	−	+	−	+	−	−	+	16	12	15	16	16	15.0	1.73
7	−	+	+	−	−	+	−	+	6	5	6	5	3	5.0	1.22
8	+	+	+	−	−	−	+	−	15	16	16	15	14	15.2	0.84
9	−	−	−	+	+	+	−	+	1	2	3	3	2	2.2	0.84
10	+	−	−	+	+	−	+	−	7	11	4	7	6	7.0	2.55
11	−	+	−	+	−	+	+	−	13	3	8	12	8	8.8	3.96
12	+	+	−	+	−	−	−	+	3	1	5	1	4	2.8	1.79
13	−	−	+	+	−	−	+	+	2	10	2	4	5	9.6	3.29
14	+	−	+	+	−	+	−	−	4	4	1	2	1	2.4	1.52
15	−	+	+	+	+	−	−	−	5	15	9	6	11	9.2	4.02
16	+	+	+	+	+	+	+	+	11	14	12	13	13	12.6	1.14

c) Usar la desviación estándar de las calificaciones (o alguna transformación apropiada tal como log *s*) como variable de respuesta. ¿Qué conclusiones pueden sacarse acerca de los efectos de las ocho variables sobre la variabilidad de la calidad del vino?

d) Después de mirar los resultados, Harry y Judy coincidieron en que uno de los miembros del panel (DCM) sabía más de cerveza que de vino, por lo que decidieron eliminar su calificación. ¿Qué efecto tendría esto en los resultados y las conclusiones de los incisos *b* y *c*?

e) Suponga que justo antes de empezar el experimento, Harry y Judy se enteraron de que las ocho nuevas barricas que ordenaron de Francia para usarlas en el experimento no llegarían a tiempo, y que las 16 corridas tendrían que hacerse con las barricas viejas. Si Harry y Judy simplemente eliminan la columna *C* de su diseño, ¿qué ocurre con las relaciones de los alias? ¿Es necesario que empiecen de nuevo y construyan otro diseño?

f) Harry y Judy saben por experiencia que es improbable que algunas de las combinaciones de tratamientos produzcan buenos resultados. Por ejemplo, la corrida con las ocho variables en el nivel alto generalmente resulta en un vino con una calificación baja. Esto se confirmó el 8 de marzo de 1986 en la prueba del vino. Quieren establecer un nuevo diseño para su Pinot Noir 1986 utilizando estas mismas ocho variables, pero no quieren correr el experimento con los ocho factores en el nivel alto. ¿Qué diseño sugeriría el lector?

8-27. En un artículo de *Quality Engineering* ("Una aplicación de los diseños experimentales factoriales fraccionados", vol. 1, pp. 19-23) M.B. Kilgo describe un experimento para determinar el efecto de la presión del CO_2 (*A*), la temperatura del CO_2 (*B*), la humedad del cacahuate (*C*), la velocidad de flujo del CO_2 (*D*) y el tamaño de las partículas de cacahuate (*E*) sobre el rendimiento total del aceite por lote de cacahuates (*y*). Los niveles que usó para estos factores son los siguientes:

Nivel codificado	*A*, presión (bar)	*B*, temperatura (°C)	*C*, humedad (% por peso)	*D*, flujo (litros/min)	*E*, tamaño de las partículas (mm)
−1	415	25	5	40	1.28
1	550	95	15	60	4.05

Kilgo realizó el experimento factorial fraccionado con 16 corridas que se muestra a continuación.

A	*B*	*C*	*D*	*E*	*y*
415	25	5	40	1.28	63
550	25	5	40	4.05	21
415	95	5	40	4.05	36
550	95	5	40	1.28	99
415	25	15	40	4.05	24
550	25	15	40	1.28	66
415	95	15	40	1.28	71
550	95	15	40	4.05	54
415	25	5	60	4.05	23
550	25	5	60	1.28	74
415	95	5	60	1.28	80
550	95	5	60	4.05	33
415	25	15	60	1.28	63
550	25	15	60	4.05	21
415	95	15	60	4.05	44
550	95	15	60	1.28	96

a) ¿Qué tipo de diseño se ha utilizado? Identificar la relación de definición y las relaciones de los alias.

b) Estimar los efectos de los factores y usar una gráfica de probabilidad normal para hacer la identificación tentativa de los factores importantes.

c) Efectuar el análisis estadístico apropiado para probar las hipótesis de que los factores identificados en el inciso *b* anterior tienen un efecto significativo sobre el rendimiento del aceite de cacahuate.

d) Ajustar un modelo que pueda usarse para predecir el rendimiento del aceite de cacahuate en términos de los factores que se han identificado como importantes.

e) Analizar los residuales de este experimento y comentar la adecuación del modelo.

8-28. Los ingenieros de la planta Essex Aluminum de Ford Motor Company llevaron a cabo un experimento factorial fraccionado en 10 factores con 16 corridas para el vaciado en arena de tubos múltiples para motor, el cual se describe en el artículo "Estudio del proceso de vaciado evaporativo para múltiples de admisión de 3.0 litros 'Poor Sandfill'", de D. Becknell (*Fourth Symposium on Taguchi Methods*, American Supplier Institute, Dearborn, MI, pp. 120-130). El objetivo fue determinar cuáles de los 10 factores tienen un efecto sobre la proporción de vaciados defectuosos. El diseño y la proporción resultante de vaciados no defectuosos $\hat{p}$ que se observaron en cada corrida se presentan enseguida. Se trata de una fracción de resolución III con generadores $E = CD, F = BD, G = BC, H = AC, J = AB$ y $K = ABC$. Suponga que el número de vaciados hechos en cada corrida del diseño es 1000.

Corrida	A	B	C	D	E	F	G	H	J	K	$\hat{p}$	$\text{arcsen}\sqrt{\hat{p}}$	Modificación de F&T
1	−	−	−	−	+	+	+	+	+	−	0.958	1.364	1.363
2	+	−	−	−	+	+	+	−	−	+	1.000	1.571	1.555
3	−	+	−	−	+	−	−	+	−	+	0.977	1.419	1.417
4	+	+	−	−	+	−	−	−	+	−	0.775	1.077	1.076
5	−	−	+	−	−	+	−	−	+	+	0.958	1.364	1.363
6	+	−	+	−	−	+	−	+	−	−	0.958	1.364	1.363
7	−	+	+	−	−	−	+	−	−	−	0.813	1.124	1.123
8	+	+	+	−	−	−	+	+	+	+	0.906	1.259	1.259
9	−	−	−	+	−	−	+	+	+	−	0.679	0.969	0.968
10	+	−	−	+	−	−	+	−	−	+	0.781	1.081	1.083
11	−	+	−	+	−	+	−	+	−	+	1.000	1.571	1.556
12	+	+	−	+	−	+	−	−	+	−	0.896	1.241	1.242
13	−	−	+	+	+	−	−	−	+	+	0.958	1.364	1.363
14	+	−	+	+	+	−	−	+	−	−	0.818	1.130	1.130
15	−	+	+	+	+	+	+	−	−	−	0.841	1.161	1.160
16	+	+	+	+	+	+	+	+	+	+	0.955	1.357	1.356

a) Encontrar la relación de definición y las relaciones de los alias de este diseño.

b) Estimar los efectos de los factores y usar una gráfica de probabilidad normal para hacer la identificación tentativa de los factores importantes.

c) Ajustar el modelo apropiado utilizando los factores identificados en el inciso *b* anterior.

d) Graficar los residuales de este modelo contra la proporción predicha de vaciados no defectuosos. Construir también una gráfica de probabilidad normal de los residuales. Comentar la adecuación de estas gráficas.

e) El lector habrá notado en el inciso *d* un indicio de que la varianza de la respuesta no es constante (considerando que la respuesta es una proporción, esto debería haberse anticipado). La tabla anterior también muestra una transformación de $\hat{p}$, arcsen de la raíz cuadrada de $\hat{p}$, que es de uso generalizado como *transformación para estabilizar la varianza* de los datos de la proporción (referirse a la discusión de las transformaciones para estabilizar la varianza del capítulo 3). Repetir los incisos *a* al *d* anteriores utilizan-

do la respuesta transformada y comentar los resultados. Específicamente, ¿son mejores ahora las gráficas de los residuales?

f) Hay una modificación de la transformación arcsen de la raíz cuadrada, propuesta por Freeman y Tukey ("Transformaciones relacionadas con ángulos y la raíz cuadrada", *Annals of Mathematical Statistics*, vol. 21, pp. 607-611) que mejora su desempeño en las colas. La modificación de F&T es:

$$[\text{arcsen}\sqrt{n\hat{p}/(n+1)} + \text{arcsen}\sqrt{(n\hat{p}+1)/(n+1)}]/2$$

Resolver de nuevo los incisos *a* al *d* utilizando esta transformación y comentar los resultados. (Para una interesante discusión y análisis de este experimento, referirse a "Análisis de experimentos factoriales con defectos o partes defectuosas como respuesta", de S. Bisgaard y H.T. Fuller, *Quality Engineering*, vol. 7, pp. 429-443.)

8-29. Un experimento factorial fraccionado en nueve factores y 16 corridas fue conducido por el departamento Chrysler Motors Engineering y se describe en el artículo "Mejoramiento del proceso compuesto de moldeo de planchas", de P.I. Hsieh y D.E. Goodwin (*Fourth Symposium on Taguchi Methods*, American Supplier Institute, Dearborn, MI, pp. 13-21). El objetivo era reducir el número de defectos en el acabado de rejillas de planchas moldeadas de recuadros abiertos. El diseño y el número resultante de defectos, *c*, observados en cada corrida se muestran a continuación. Se trata de una fracción de resolución III con generadores $E = BD$, $F = BCD$, $G = AC$, $H = ACD$ y $J = AB$.

Corrida	A	B	C	D	E	F	G	H	J	c	$\sqrt{c}$	Modificación de F&T
1	−	−	−	−	+	−	+	−	+	56	7.48	7.52
2	+	−	−	−	+	−	−	+	−	17	4.12	4.18
3	−	+	−	−	−	+	+	−	−	2	1.41	1.57
4	+	+	−	−	−	+	−	+	+	4	2.00	2.12
5	−	−	+	−	+	+	−	+	+	3	1.73	1.87
6	+	−	+	−	+	+	+	−	−	4	2.00	2.12
7	−	+	+	−	−	−	−	+	−	50	7.07	7.12
8	+	+	+	−	−	−	+	−	+	2	1.41	1.57
9	−	−	−	+	−	+	+	+	+	1	1.00	1.21
10	+	−	−	+	−	+	−	−	−	0	0.00	0.50
11	−	+	−	+	+	−	+	+	−	3	1.73	1.87
12	+	+	−	+	+	−	−	−	+	12	3.46	3.54
13	−	−	+	+	−	−	−	−	+	3	1.73	1.87
14	+	−	+	+	−	−	+	+	−	4	2.00	2.12
15	−	+	+	+	+	+	−	−	−	0	0.00	0.50
16	+	+	+	+	+	+	+	+	+	0	0.00	0.50

a) Encontrar la relación de definición y las relaciones de los alias de este diseño.

b) Estimar los efectos de los factores y usar una gráfica de probabilidad normal para hacer la identificación tentativa de los factores importantes.

c) Ajustar un modelo apropiado utilizando los factores identificados en el inciso *b* anterior.

d) Graficar los residuales de este modelo contra el número predicho de defectos. Asimismo, construir una gráfica de probabilidad normal de los residuales. Comentar la adecuación de estas gráficas.

e) El lector habrá notado en el inciso *d* un indicio de que la varianza de la respuesta no es constante (considerando que la respuesta es un conteo, esto debería haberse anticipado). La tabla anterior también incluye una transformación de *c*, la raíz cuadrada, que es una *transformación para estabilizar la varianza* de uso generalizado con datos de conteos (referirse a la exposición de las transformaciones para estabilizar la varianza del capítulo 3). Repetir los incisos *a* al *d* utilizando la respuesta transformada y comentar los resultados. Específicamente, ¿han mejorado ahora las gráficas de los residuales?

f) Hay una modificación de la transformación de la raíz cuadrada, propuesta por Freeman y Tukey ("Transformaciones relacionadas con ángulos y raíz cuadrada", *Annals of Mathematical Statistics*, vol. 21, pp. 607-611) que mejora su desempeño. La modificación de F&T de la transformación de la raíz cuadrada es:

$$[\sqrt{c}+\sqrt{(c+1)}]/2$$

Resolver de nuevo los incisos *a* al *d* utilizando esta transformación y comentar los resultados. (Para una interesante discusión y análisis de este experimento, referirse a "Análisis de experimentos factoriales con defectos o partes defectuosas como respuesta", de S. Bisgaard y H.T. Fuller, *Quality Engineering*, vol. 7, pp. 429-443.)

8-30. Se corre un experimento en una fábrica de semiconductores para investigar el efecto de seis factores sobre la amplificación del transistor. El diseño seleccionado es el 2_{IV}^{6-2} que se muestra a continuación:

Orden estándar	Orden de las corridas	*A*	*B*	*C*	*D*	*E*	*F*	Amplificación
1	2	−	−	−	−	−	−	1455
2	8	+	−	−	−	+	−	1511
3	5	−	+	−	−	+	+	1487
4	9	+	+	−	−	−	+	1596
5	3	−	−	+	−	+	+	1430
6	14	+	−	+	−	−	+	1481
7	11	−	+	+	−	−	−	1458
8	10	+	+	+	−	+	−	1549
9	15	−	−	−	+	−	+	1454
10	13	+	−	−	+	+	+	1517
11	1	−	+	−	+	+	−	1487
12	6	+	+	−	+	−	−	1596
13	12	−	−	+	+	+	−	1446
14	4	+	−	+	+	−	−	1473
15	7	−	+	+	+	−	+	1461
16	16	+	+	+	+	+	+	1563

a) Usar una gráfica normal de los efectos para identificar los factores significativos.

b) Conducir las pruebas estadísticas apropiadas para el modelo identificado en el inciso *a*.

c) Analizar los residuales y comentar los resultados.

d) ¿Es posible encontrar un conjunto de condiciones de operación que produzca una amplificación de 1500 ± 25?

8-31. El tratamiento térmico es de uso común para carbonizar piezas metálicas, como engranes. El espesor de la capa carbonizada es una variable de salida crítica de este proceso, y suele medirse realizando un análisis de carbono del paso del engrane (la cara superior del diente del engrane). Se estudiaron seis factores en un diseño 2_{IV}^{6-2}: *A* = temperatura del horno, *B* = duración del ciclo, *C* = concentración de carbono, *D* = duración del

ciclo de carbonización, E = concentración de carbono del ciclo difuso y F = duración del ciclo difuso. El experimento se presenta a continuación:

Orden estándar	Orden de las corridas	A	B	C	D	E	F	Paso
1	5	−	−	−	−	−	−	74
2	7	+	−	−	−	+	−	190
3	8	−	+	−	−	+	+	133
4	2	+	+	−	−	−	+	127
5	10	−	−	+	−	+	+	115
6	12	+	−	+	−	−	+	101
7	16	−	+	+	−	−	−	54
8	1	+	+	+	−	+	−	144
9	6	−	−	−	+	−	+	121
10	9	+	−	−	+	+	+	188
11	14	−	+	−	+	+	−	135
12	13	+	+	−	+	−	−	170
13	11	−	−	+	+	+	−	126
14	3	+	−	+	+	−	−	175
15	15	−	+	+	+	−	+	126
16	4	+	+	+	+	+	+	193

a) Estimar los efectos de los factores y representarlos en una gráfica de probabilidad normal. Seleccionar un modelo tentativo.

b) Efectuar las pruebas estadísticas apropiadas en el modelo.

c) Analizar los residuales y comentar la adecuación del modelo.

d) Interpretar los resultados de este experimento. Suponer que es deseable un espesor de la capa de entre 140 y 160.

8-32. Se estudian cinco factores en el diseño factorial fraccionado irregular de resolución V mostrado enseguida:

Orden estándar	Orden de las corridas	A	B	C	D	E	y
1	1	−	−	−	−	−	16.33
2	10	−	+	−	−	−	18.43
3	5	+	+	−	−	−	27.07
4	4	−	−	+	−	−	16.95
5	15	+	−	+	−	−	14.58
6	19	−	+	+	−	−	19.12
7	16	−	−	−	+	−	18.96
8	7	+	−	−	+	−	23.56
9	8	+	+	−	+	−	29.15
10	3	+	−	+	+	−	15.74
11	13	−	+	+	+	−	20.73
12	11	+	+	+	+	−	21.52
13	12	−	−	−	−	+	15.58
14	20	+	−	−	−	+	21.03
15	9	+	+	−	−	+	26.78

Orden estándar	Orden de las corridas	A	B	C	D	E	y
16	22	+	–	+	–	+	13.39
17	21	–	+	+	–	+	18.63
18	6	+	+	+	–	+	19.01
19	23	–	–	–	+	+	17.96
20	18	–	+	–	+	+	20.49
21	24	+	+	–	+	+	29.31
22	17	–	–	+	+	+	17.62
23	2	+	–	+	+	+	16.03
24	14	–	+	+	+	+	21.42

a) Analizar los datos de este experimento. ¿Qué factores influyen en la respuesta y?
b) Analizar los residuales. Comentar la adecuación del modelo.

9 Diseños factoriales y factoriales fraccionados con tres niveles y con niveles mixtos

Las series con dos niveles de los diseños factoriales y factoriales fraccionados que se comentaron en los capítulos 6, 7 y 8 son de uso generalizado en la investigación y el desarrollo industrial. Hay algunas extensiones y variantes de estos diseños que en ocasiones son útiles, como los diseños para los casos en que todos los factores están presentes con tres niveles. Estos diseños 3^k se analizan en este capítulo. Se consideran también los casos en que algunos de los factores tienen dos niveles y otros factores tienen ya sea tres o cuatro niveles.

9-1 DISEÑO FACTORIAL 3^k

9-1.1 Notación y motivación del diseño 3^k

Se estudia ahora el diseño factorial 3^k; es decir, un arreglo factorial de k factores que tienen tres niveles cada uno. Se usarán letras mayúsculas para denotar los factores y las interacciones. Se hará referencia a los tres niveles de los factores como bajo, intermedio y alto. Hay varias notaciones diferentes que se usan para representar estos niveles de los factores; una posibilidad es representar los niveles de los factores con los dígitos 0 (bajo), 1 (intermedio) y 2 (alto). Cada combinación de tratamientos del diseño 3^k se denotará por k dígitos, donde el primer dígito indica el nivel del factor A, el segundo dígito indica el nivel del factor B, ..., y el dígito k-ésimo indica el nivel del factor K. Por ejemplo, en un diseño 3^2, 00 denota la combinación de tratamientos correspondiente a A y B ambos en el nivel bajo, y 01 denota la combinación de tratamientos correspondiente a A en el nivel bajo y B en el nivel intermedio. En las figuras 9-1 y 9-2 se muestra la representación geométrica de los diseños 3^2 y 3^3, respectivamente, utilizando esta notación.

Este sistema de notación pudo haberse usado en los diseños 2^k presentados anteriormente, utilizando 0 y 1 en lugar del 1 negativo y el 1 positivo, respectivamente. En el diseño 2^k se prefirió la notación ± 1 porque facilita la vista geométrica del diseño y porque puede aplicarse directamente al modelado de regresión, la separación en bloques y la construcción de factoriales fraccionados.

En el sistema de los diseños 3^k, cuando los factores son **cuantitativos,** es común denotar los niveles bajo, intermedio y alto con -1, 0 y $+1$, respectivamente. Con esto se facilita el ajuste de un **modelo de re-**

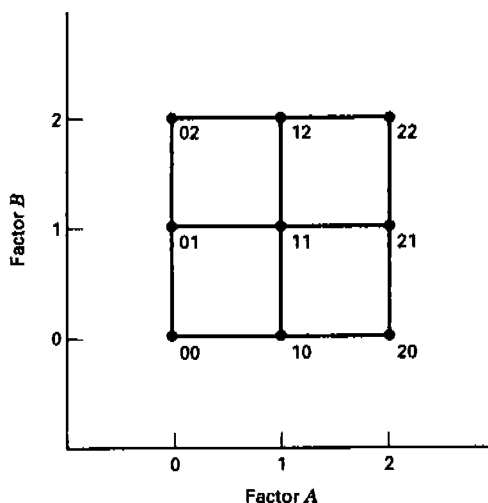


Figura 9-1 Combinaciones de tratamientos en un diseño 3^2.

gresión que relaciona la respuesta con los niveles de los factores. Por ejemplo, considere el diseño 3^2 de la figura 9-1, y sea que x_1 represente al factor A y que x_2 represente al factor B. Un modelo de regresión que relaciona la respuesta y con x_1 y x_2 que se basa en este diseño es

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \varepsilon \tag{9-1}$$

Observe que la adición de un tercer nivel de los factores permite que la relación entre la respuesta y los factores del diseño se modele como un modelo cuadrático.

El diseño 3^k es ciertamente una elección posible para un experimentador que se preocupa por la curvatura en la función de respuesta. Sin embargo, es necesario tomar en consideración dos puntos:

1. El diseño 3^k no es la forma más eficiente de modelar una relación cuadrática; los **diseños de superficie de respuesta** que se exponen en el capítulo 11 son alternativas superiores.

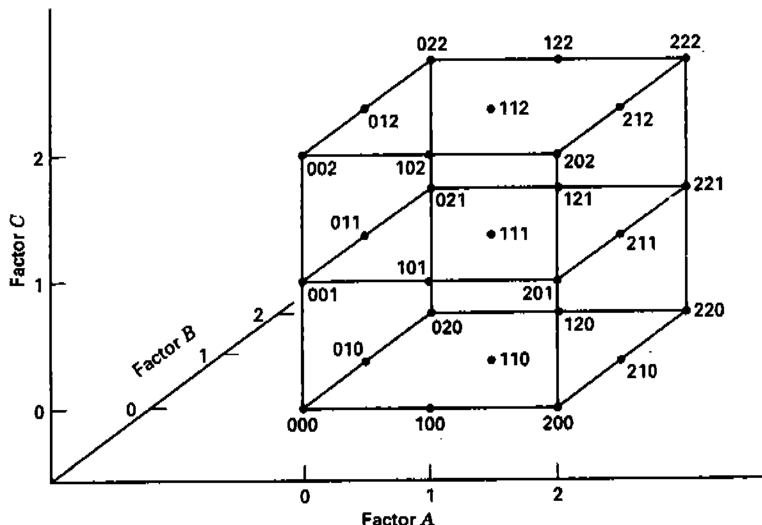


Figura 9-2 Combinaciones de tratamientos en un diseño 3^3.

2. El diseño 2^k aumentado con los puntos centrales, como se analizó en el capítulo 6, es una forma excelente de obtener una indicación de la curvatura. Permite conservar reducido el tamaño y la complejidad del diseño y al mismo tiempo permite obtener cierta protección contra la curvatura. Entonces, si la curvatura es importante, el diseño de dos niveles puede aumentarse con **corridas axiales** para obtener un **diseño central compuesto,** como se ilustra en la figura 6-36. Esta **estrategia secuencial** de experimentación es más eficiente, por mucho, que correr un diseño factorial 3^k con factores cuantitativos.

9-1.2 El diseño 3^2

El diseño más simple del sistema 3^k es el diseño 3^2, el cual tiene dos factores, cada uno con tres niveles. Las combinaciones de tratamientos de este diseño se mostraron en la figura 9-1. Puesto que están presentes $3^2 = 9$ combinaciones de tratamientos, hay ocho grados de libertad entre estas combinaciones de tratamientos. Los efectos principales de A y B tienen dos grados de libertad cada uno, y la interacción AB tiene cuatro grados de libertad. Si hay n réplicas, habrá $n3^2 - 1$ grados de libertad totales y $3^2(n-1)$ grados de libertad del error.

Las sumas de cuadrados de A, B y AB pueden calcularse mediante los métodos usuales para los diseños factoriales analizados en el capítulo 5. Cada efecto principal puede representarse con un componente lineal y uno cuadrático, cada uno con un solo grado de libertad, como se observa en la ecuación 9-1. Desde luego, esto sólo tiene sentido si el factor es cuantitativo.

La partición de la interacción de dos factores AB puede hacerse de dos maneras. El primer método consiste en subdividir AB en los cuatro componentes con un solo grado de libertad que corresponden a $AB_{L\times L}$, $AB_{L\times Q}$, $AB_{Q\times L}$ y $AB_{Q\times Q}$. Esto puede hacerse ajustando los términos $\beta_{12}x_1x_2$, $\beta_{122}x_1x_2^2$, $\beta_{112}x_1^2x_2$ y $\beta_{1122}x_1^2x_2^2$, respectivamente, como se indicó en el ejemplo 5-5. Para los datos de la vida de la herramienta de ese ejemplo se obtiene $SS_{AB_{L\times L}} = 8.00$, $SS_{AB_{L\times Q}} = 42.67$, $SS_{AB_{Q\times L}} = 2.67$ y $SS_{AB_{Q\times Q}} = 8.00$. Puesto que se trata de una partición ortogonal de AB, observe que $SS_{AB} = SS_{AB_{L\times L}} + SS_{AB_{L\times Q}} + SS_{AB_{Q\times L}} + SS_{AB_{Q\times Q}} = 61.34$.

El segundo método se basa en los **cuadrados latinos ortogonales.** Considere los totales de las combinaciones de los tratamientos para los datos del ejemplo 5-5. Estos totales se muestran en la figura 9-3 como los números encerrados en círculos dentro de los cuadrados. Los dos factores A y B corresponden a los renglones y las columnas, respectivamente, de un cuadrado latino 3×3. En la figura 9-3 se muestran dos cuadrados latinos 3×3 particulares, superpuestos en los totales de las celdas.

Estos dos cuadrados latinos son **ortogonales;** es decir, si uno de los cuadrados se superpone en el otro, cada letra del primer cuadrado aparecerá exactamente una vez con cada letra del segundo cuadrado. Los totales de las letras en el cuadrado a son $Q = 18$, $R = -2$ y $S = 8$, y la suma de cuadrados entre estos totales

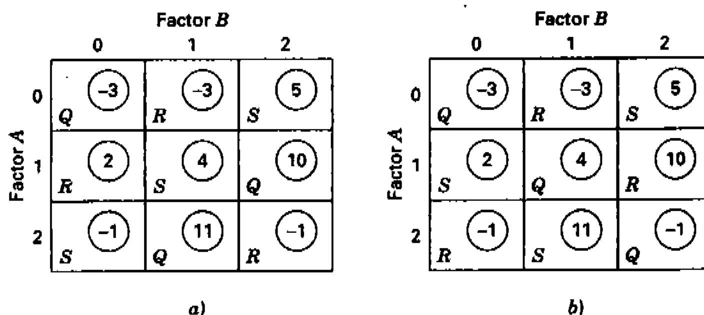


Figura 9-3 Totales de las combinaciones de los tratamientos del ejemplo 5-5 con dos cuadrados latinos superpuestos.

es $[18^2 + (-2)^2 + 8^2]/(3)(2) - [24^2/(9)(2)] = 33.34$, con dos grados de libertad. De manera similar, los totales de las letras en el cuadrado b son $Q = 0, R = 6$ y $S = 18$, y la suma de cuadrados entre estos totales es $[0^2 + 6^2 + 18^2]/(3)(2) - [24^2/(9)(2)] = 28.00$, con dos grados de libertad. Observe que la suma de estos dos componentes es

$$33.34 + 28.00 = 61.34 = SS_{AB}$$

con $2 + 2 = 4$ grados de libertad.

En general, a la suma de cuadrados calculada con el cuadrado a se le llama el **componente AB de la interacción,** y a la suma de cuadrados calculada con el cuadrado b se le llama el **componente AB^2 de la interacción.** Cada uno de los componentes AB y AB^2 tiene dos grados de libertad. Se usa esta terminología porque si los niveles (0, 1, 2) de A y B se denotan por x_1 y x_2, respectivamente, entonces se encuentra que las letras ocupan celdas de acuerdo con el siguiente patrón:

Cuadrado *a*	**Cuadrado *b***
$Q: x_1 + x_2 = 0 \pmod 3$	$Q: x_1 + 2x_2 = 0 \pmod 3$
$R: x_1 + x_2 = 1 \pmod 3$	$S: x_1 + 2x_2 = 1 \pmod 3$
$S: x_1 + x_2 = 2 \pmod 3$	$R: x_1 + 2x_2 = 2 \pmod 3$

Por ejemplo, en el cuadrado b se observa que la celda de en medio corresponde a $x_1 = 1$ y $x_2 = 1$; por lo tanto, $x_1 + 2x_2 = 1 + (2)(1) = 3 = 0 \pmod 3$, y Q ocuparía la celda de en medio. Cuando se consideran expresiones de la forma A^pB^q, se establece la convención de que el único exponente permitido en la primera letra es 1. Si el exponente de la primera letra no es 1, la expresión completa se eleva al cuadrado y los exponentes se reducen al módulo 3. Por ejemplo, A^2B es lo mismo que AB^2 porque

$$A^2B = (A^2B)^2 = A^4B^2 = AB^2$$

Los componentes AB y AB^2 de la interacción AB no tienen significado real y por lo general no se incluyen en la tabla del análisis de varianza. Sin embargo, esta partición en gran medida arbitraria de la interacción AB en dos componentes ortogonales con dos grados de libertad es muy útil para construir diseños más complicados. Además, no hay relación entre los componentes AB y AB^2 de la interacción y las sumas de cuadrados de $AB_{L \times L}$, $AB_{L \times Q}$, $AB_{Q \times L}$ y $AB_{Q \times Q}$.

Los componentes AB y AB^2 de la interacción pueden calcularse de otra manera. Considere los totales de las combinaciones de los tratamientos en cualquiera de los cuadrados de la figura 9-3. Si se hace la suma de los datos en las diagonales hacia abajo de izquierda a derecha, se obtienen los totales $-3 + 4 - 1 = 0$, $-3 + 10 - 1 = 6$ y $5 + 11 + 2 = 18$. La suma de cuadrados entre estos totales es 28.00 (AB^2). En forma similar, los totales de la diagonal hacia abajo de derecha a izquierda son $5 + 4 - 1 = 8$, $-3 + 2 - 1 = -2$ y $-3 + 11 + 10 = 18$. La suma de cuadrados entre estos totales es 33.34 (AB). Yates llamó a estos componentes de la interacción los **componentes I y J de la interacción,** respectivamente. Se usarán aquí indistintamente las dos notaciones; es decir,

$$I(AB) = AB^2$$
$$J(AB) = AB$$

9-1.3 El diseño 3^3

Suponga ahora que hay tres factores (*A*, *B* y *C*) bajo estudio, y que cada factor tiene tres niveles dispuestos en un experimento factorial. Se trata de un diseño factorial 3^3, y la disposición experimental y la notación de las combinaciones de los tratamientos se presentaron anteriormente en la figura 9-2. Las 27 combinaciones de tratamientos tienen 26 grados de libertad. Cada efecto principal tiene 2 grados de libertad, cada interacción de dos factores tiene 4 grados de libertad y la interacción de tres factores tiene 8 grados de libertad. Si se hacen *n* réplicas, hay $n3^3 - 1$ grados de libertad totales y $3^3(n-1)$ grados de libertad del error.

Las sumas de cuadrados pueden calcularse utilizando los métodos estándares para los diseños factoriales. Además, si los factores son cuantitativos, es posible hacer la partición de los efectos principales en un componente lineal y uno cuadrático, cada uno con un solo grado de libertad. Las interacciones de dos factores pueden descomponerse en efectos lineal × lineal, lineal × cuadrático, cuadrático × lineal y cuadrático × cuadrático. Por último, puede hacerse la partición de la interacción de tres factores *ABC* en ocho componentes con un solo grado de libertad que corresponden a lineal × lineal × lineal, lineal × lineal × cuadrático, etcétera. Esta descomposición de la interacción de tres factores no es por lo general de gran utilidad.

También es posible hacer la partición de las interacciones de dos factores en sus componentes *I* y *J*. Éstos se designarían AB, AB^2, AC, AC^2, BC y BC^2, y cada componente tendría dos grados de libertad. Como en el diseño 3^2, estos componentes no tienen significación física.

Es posible hacer la partición de la interacción de tres factores *ABC* en cuatro componentes ortogonales con dos grados de libertad, a los que suele denominarse los componentes *W*, *X*, *Y* y *Z* de la interacción. También se hace referencia a ellos como los componentes AB^2C^2, AB^2C, ABC^2 y ABC de la interacción *ABC*, respectivamente. Las dos notaciones se usan indistintamente; es decir,

$$W(ABC) = AB^2C^2$$
$$X(ABC) = AB^2C$$
$$Y(ABC) = ABC^2$$
$$Z(ABC) = ABC$$

Observe que ninguna de las primeras letras puede tener un exponente diferente de 1. Al igual que los componentes *I* y *J*, los componentes *W*, *X*, *Y* y *Z* no tienen ninguna interpretación práctica. Sin embargo, son útiles para construir diseños más complejos.

EJEMPLO 9-1 ....

Se usa una máquina para llenar contenedores metálicos de 5 galones con jarabe para una bebida gaseosa. La variable de interés es la cantidad de jarabe perdida debido al espumeo. Se piensa que tres factores influyen en el espumeo: el diseño de la boquilla (*A*), la velocidad del llenado (*B*) y la presión de operación (*C*). Se seleccionan tres boquillas, tres velocidades de llenado y tres presiones, y se corren dos réplicas de un experimento factorial 3^3. En la tabla 9-1 se muestran los datos codificados.

El análisis de varianza de los datos de la pérdida de jarabe se muestra en la tabla 9-2: Las sumas de cuadrados se calcularon con los métodos usuales. Se observa que la velocidad de llenado y la presión de operación son estadísticamente significativas. Las tres interacciones de dos factores también son significativas. En la figura 9-4 se analizan gráficamente las interacciones de dos factores. El nivel intermedio de la velocidad produce el mejor desempeño, mientras que las boquillas tipo 2 y 3, y la presión baja (10 psi) o bien alta (20 psi) parecen ser las más efectivas para reducir la pérdida de jarabe.

....

Tabla 9-1 Datos de la pérdida de jarabe del ejemplo 9-1 (las unidades son centímetros cúbicos – 70)

	Tipo de boquilla (*A*)								
	1			2			3		
	Velocidad (en rpm) (*B*)								
Presión (en psi) (*C*)	100	120	140	100	120	140	100	120	140
10	–35	–45	–40	17	–65	20	–39	–55	15
	–25	–60	15	24	–58	4	–35	–67	–30
15	110	–10	80	55	–55	110	90	–28	110
	75	30	54	120	–44	44	113	–26	135
20	4	–40	31	–23	–64	–20	–30	–61	54
	5	–30	36	–5	–62	–31	–55	–52	4

El ejemplo 9-1 ilustra una situación en la que el diseño de tres niveles suele encontrar cierta aplicación; uno o más de los factores es **cualitativo,** asumiendo desde luego tres niveles, y los demás factores son **cuantitativos.** En este ejemplo, suponga que sólo hay tres diseños de la boquilla que son de interés. Se trata evidentemente, entonces, de un factor cualitativo que requiere tres niveles. La velocidad de llenado y la presión de operación son factores cuantitativos. Por lo tanto, podría ajustarse un modelo cuadrático como el de la ecuación 9-1 en los dos factores, velocidad y presión, con cada nivel del factor boquilla.

En la tabla 9-3 se muestran estos modelos de regresión cuadráticos. Las β de estos modelos se estimaron usando un programa de computadora de regresión lineal estándar. (En el capítulo 10 se analizará con mayor detalle la regresión de mínimos cuadrados.) En estos modelos, las variables x_1 y x_2 están codificadas en los niveles –1, 0, +1, como se estudió anteriormente, y se supusieron los siguientes niveles naturales para la presión y la velocidad:

Nivel codificado	Velocidad (rpm)	Presión (psi)
–1	100	10
0	120	15
+1	140	20

En la tabla 9-3 se presentan estos modelos tanto en términos de estas variables codificadas como en términos de los niveles naturales de la velocidad y la presión.

Tabla 9-2 Análisis de varianza de los datos de la pérdida de jarabe

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0	Valor *P*
A, boquilla	993.77	2	496.89	1.17	0.3256
B, velocidad	61,190.33	2	30,595.17	71.74	<0.0001
C, presión	69,105.33	2	34,552.67	81.01	<0.0001
AB	6,300.90	4	1,575.22	3.69	0.0383
AC	7,513.90	4	1,878.47	4.40	0.0222
BC	12,854.34	4	3,213.58	7.53	0.0025
ABC	4,628.76	8	578.60	1.36	0.2737
Error	11,515.50	27	426.50		
Total	174,102.83	53			

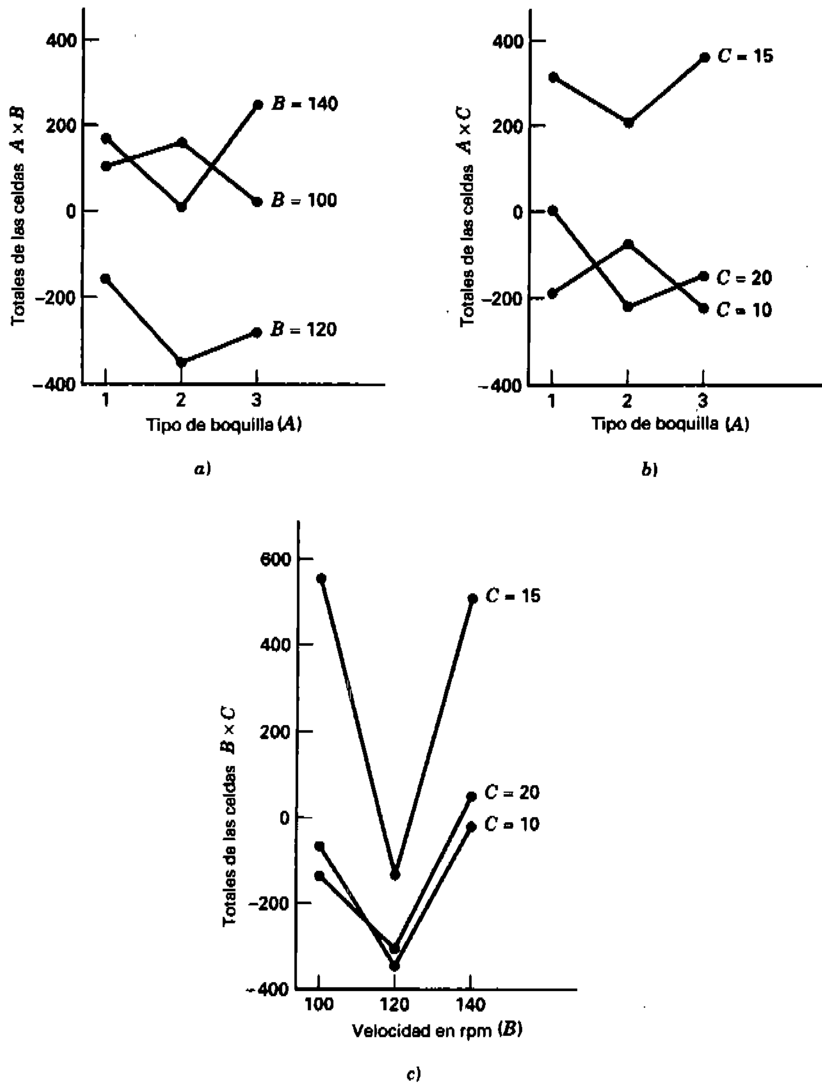


Figura 9-4 Interacciones de dos factores del ejemplo 9-1.

Tabla 9-3 Modelos de regresión para el ejemplo 9-1

Tipo de boquilla	x_1 = velocidad (S), x_2 = Presión (P) en unidades codificadas
1	$\hat{y} = 22.1 + 3.5x_1 + 16.3x_2 + 51.7x_1^2 - 71.8x_2^2 + 2.9x_1x_2$
	$\hat{y} = 1217.3 - 31.256S + 86.017P + 0.12917S^2 - 2.8733P^2 + 0.02875SP$
2	$\hat{y} = 25.6 - 22.8x_1 - 12.3x_2 + 14.1x_1^2 - 56.9x_2^2 - 0.7x_1x_2$
	$\hat{y} = 180.1 - 9.475S + 66.75P + 0.035S^2 - 2.2767P^2 - 0.0075SP$
3	$\hat{y} = 15.1 + 20.3x_1 + 5.9x_2 + 75.8x_1^2 - 94.9x_2^2 + 10.5x_1x_2$
	$\hat{y} = 1940.1 - 40.058S + 102.48P + 0.18958S^2 - 3.7967P^2 + 0.105SP$

En la figura 9-5 se muestran las gráficas de contorno de las superficies de respuesta de la pérdida de jarabe constante, como una función de la velocidad y la presión para cada tipo de boquilla. Estas gráficas revelan información de considerable utilidad acerca del desempeño de este sistema de llenado. Puesto que el objetivo es minimizar la pérdida de jarabe, se preferiría la boquilla tipo 3, ya que los contornos observados más pequeños (–60) sólo aparecen en esta gráfica. Deberán usarse la velocidad de llenado cerca del nivel intermedio de 120 rpm y el nivel de presión ya sea bajo o alto.

Cuando se construyen gráficas de contorno para un experimento que tiene una mezcla de factores cuantitativos y cualitativos, no es raro encontrar que las formas de las superficies de respuesta de los factores cuantitativos son muy diferentes en cada nivel de los factores cualitativos. Esto puede observarse en cierta medida en la figura 9-5, donde la forma de la superficie para la boquilla tipo 2 es considerablemente alargada en comparación con las superficies de las boquillas tipo 1 y 3. Cuando esto ocurre, implica que las condiciones de operación óptimas (y otras conclusiones importantes) en términos de los factores cuantitativos son muy diferentes en cada nivel de los factores cualitativos.

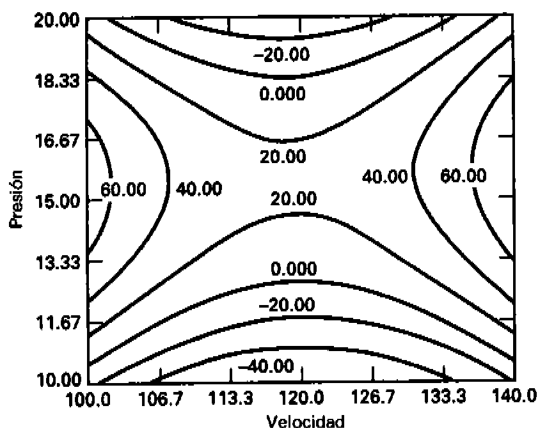
Es sencillo mostrar la partición numérica de la interacción *ABC* en sus cuatro componentes ortogonales con dos grados de libertad utilizando los datos del ejemplo 9-1. El procedimiento general ha sido descrito por Cochran y Cox [26] y Davies [36]. Primero se seleccionan dos cualesquiera de los tres factores, por ejemplo *AB*, y se calculan los totales *I* y *J* de la interacción *AB* en cada nivel del tercer factor *C*. Estos cálculos se presentan a continuación:

		A			Totales	
C	*B*	1	2	3	*I*	*J*
	100	–60	41	–74	–198	–222
10	120	–105	–123	–122	–106	–79
	140	–25	24	–15	–155	–158
	100	185	175	203	331	238
15	120	20	–99	–54	255	440
	140	134	154	245	377	285
	100	9	–28	–85	–59	–144
20	120	–70	–126	–113	–74	–40
	140	67	–51	58	–206	–155

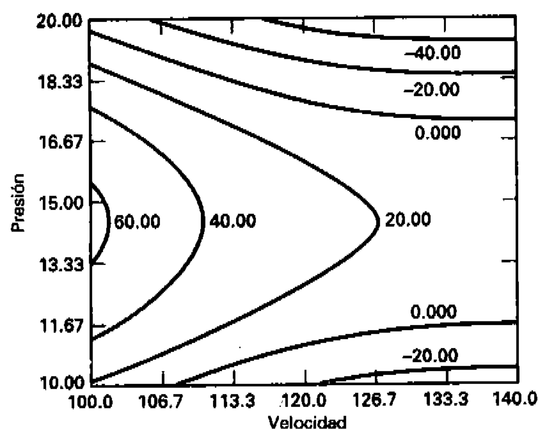
Después, los totales *I*(*AB*) y *J*(*AB*) se arreglan en una tabla de dos vías con el factor *C*, y se calculan los totales de las diagonales *I* y *J* de esta nueva disposición:

				Totales						Totales	
C		*I*(*AB*)		*I*	*J*	*C*		*J*(*AB*)		*I*	*J*
10	–198	–106	–155	–149	41	10	–222	–79	–158	63	138
15	331	255	377	212	19	15	238	440	285	62	4
20	–59	–74	–206	102	105	20	–144	–40	–155	40	23

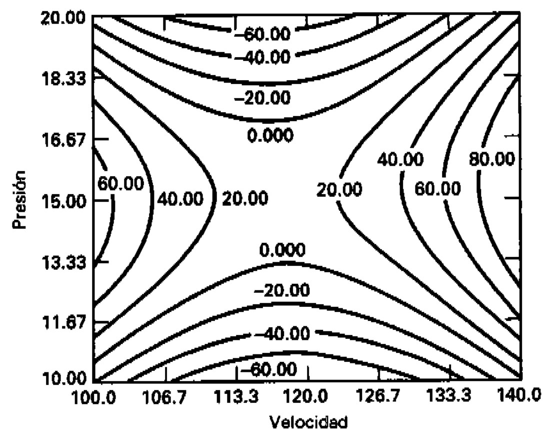
Los totales de las diagonales *I* y *J* calculados arriba son en realidad los totales que representan las cantidades $I[I(AB) \times C] = AB^2C^2$, $J[I(AB) \times C] = AB^2C$, $I[J(AB) \times C] = ABC^2$ y $J[J(AB) \times C] = ABC$, o los componentes *W*, *X*, *Y* y *Z* de *ABC*. Las sumas de cuadrados se encuentran de la manera usual; es decir,



a) Boquilla tipo 1



b) Boquilla tipo 2



c) Boquilla tipo 3

Figura 9-5 Contornos de la pérdida de jarabe constante (unidades: cc – 70) como una función de la velocidad y la presión para las boquillas tipo 1, 2 y 3, ejemplo 9-1.

$$I[I(AB)\times C] = AB^2C^2 = W(ABC)$$
$$= \frac{(-149)^2+(212)^2+(102)^2}{18} - \frac{(165)^2}{54} = 3804.11$$
$$J[I(AB)\times C] = AB^2C = X(ABC)$$
$$= \frac{(41)^2+(19)^2+(105)^2}{18} - \frac{(165)^2}{54} = 221.77$$
$$I[J(AB)\times C] = ABC^2 = Y(ABC)$$
$$= \frac{(63)^2+(62)^2+(40)^2}{18} - \frac{(165)^2}{54} = 18.77$$
$$J[J(AB)\times C] = ABC = Z(ABC)$$
$$= \frac{(138)^2+(4)^2+(23)^2}{18} - \frac{(165)^2}{54} = 584.11$$

Aun cuando se trata de una partición ortogonal de SS_{ABC}, se señala de nuevo que no se acostumbra presentarla en la tabla del análisis de varianza. En secciones subsecuentes se analiza la necesidad ocasional de calcular uno o más de estos componentes.

9.1.4 El diseño general 3^k

Los conceptos utilizados en los diseños 3^2 y 3^3 pueden extenderse de inmediato al caso de k factores, cada uno con tres niveles, es decir, a un diseño factorial 3^k. Se emplea la notación digital usual para las combinaciones de tratamientos, por lo que 0120 representa una combinación de tratamientos en un diseño 3^4 con A y D en los niveles bajos, B en el nivel intermedio y C en el nivel alto. Hay 3^k combinaciones de tratamientos, con $3^k - 1$ grados de libertad entre ellas. Estas combinaciones de tratamientos permiten determinar las sumas de cuadrados de k efectos principales, cada uno con dos grados de libertad; $\binom{k}{2}$ interacciones de dos factores, cada una con cuatro grados de libertad; ...; y una interacción de k factores con 2^k grados de libertad. En general, una interacción de h factores tiene 2^h grados de libertad. Si se hacen n réplicas, hay $n3^k - 1$ grados de libertad totales y $3^k(n-1)$ grados de libertad del error.

Las sumas de cuadrados de los efectos y las interacciones se calculan con los métodos usuales para los diseños factoriales. De manera típica, no se hace ninguna descomposición adicional de las interacciones de tres factores y de órdenes superiores. Sin embargo, cualquier interacción de h factores tiene 2^{h-1} componentes ortogonales con dos grados de libertad. Por ejemplo, la interacción de cuatro factores $ABCD$ tiene $2^{4-1} = 8$ componentes ortogonales con dos grados de libertad, denotados por $ABCD^2$, ABC^2D, AB^2CD, $ABCD$, ABC^2D^2, AB^2C^2D, AB^2CD^2 y $AB^2C^2D^2$. Al escribirse estos componentes, observe que el único exponente permitido en la primera letra es 1. Si el exponente de la primera letra no es 1, entonces la expresión completa debe elevarse al cuadrado y los exponentes deben reducirse al módulo 3. Para ilustrar lo anterior, considere

$$A^2BCD = (A^2BCD)^2 = A^4B^2C^2D^2 = AB^2C^2D^2$$

Estos componentes de la interacción no tienen ninguna interpretación física, pero son útiles para construir diseños más complejos.

El tamaño del diseño se incrementa rápidamente con k. Por ejemplo, un diseño 3^3 tiene 27 combinaciones de tratamientos por réplica, un diseño 3^4 tiene 81, un diseño 3^5 tiene 243, etcétera. Por lo tanto, con frecuencia sólo se considera una sola réplica del diseño 3^k, y las interacciones de órdenes superiores se combinan para proporcionar una estimación del error. Como una ilustración, si las interacciones de tres

factores y de órdenes superiores son insignificantes, entonces una sola réplica del diseño 3^3 proporciona 8 grados de libertad del error, y una sola réplica del diseño 3^4 proporciona 48 grados de libertad del error. Estos diseños siguen siendo grandes para $k \geq 3$ y, por consiguiente, son de escasa utilidad.

9-2 CONFUSIÓN EN EL DISEÑO FACTORIAL 3^k

Incluso cuando se considera una sola réplica del diseño 3^k, ésta requiere tantas corridas que es improbable que puedan hacerse las 3^k corridas bajo condiciones uniformes. Por lo tanto, con frecuencia es necesario hacer la confusión (o mezclado) en bloques. El diseño 3^k puede confundirse en 3^p bloques incompletos, donde $p < k$. Por lo tanto, estos diseños pueden confundirse en tres bloques, nueve bloques, etcétera.

9-2.1 El diseño factorial 3^k en tres bloques

Suponga que se quiere confundir el diseño 3^k en tres bloques incompletos. Estos tres bloques tienen dos grados de libertad entre ellos; por lo tanto, debe haber dos grados de libertad confundidos con los bloques. Recuerde que en la serie factorial 3^k cada efecto principal tiene dos grados de libertad. Además, cada interacción de dos factores tiene cuatro grados de libertad y puede descomponerse en dos componentes de la interacción (por ejemplo, AB y AB^2), cada uno con dos grados de libertad; cada interacción de tres factores tiene ocho grados de libertad y puede descomponerse en cuatro componentes de la interacción (por ejemplo, ABC, ABC^2, AB^2C y AB^2C^2), cada uno con dos grados de libertad; y así sucesivamente. Por lo tanto, es conveniente confundir un componente de interacción con los bloques.

El procedimiento general consiste en construir una **definición de contrastes**

$$L = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_k x_k \tag{9-2}$$

donde α_i representa el exponente del factor i-ésimo en el efecto que va a confundirse y x_i es el nivel del factor i-ésimo en una combinación de tratamientos particular. Para la serie 3^k se tiene $\alpha_i = 0, 1$ o 2, donde la primera α_i diferente de cero es la unidad, y $x_i = 0$ (nivel bajo), 1 (nivel intermedio) o 2 (nivel alto). Las combinaciones de tratamientos del diseño 3^k se asignan a los bloques con base en el valor de L (mod 3). Puesto que L (mod 3) sólo puede asumir los valores 0, 1 o 2, tres bloques están definidos de manera única. Las combinaciones de tratamientos que satisfacen $L = 0$ (mod 3) constituyen el **bloque principal.** Este bloque incluirá siempre la combinación de tratamientos 00 ...0.

Por ejemplo, suponga que quiere construirse un diseño factorial 3^2 en tres bloques. Cualquiera de los componentes de la interacción AB, AB o AB^2, puede confundirse con los bloques. Al elegir arbitrariamente AB^2, se obtiene la definición de contrastes

$$L = x_1 + 2x_2$$

El valor de L (mod 3) de cada combinación de tratamientos puede encontrarse de la siguiente manera:

00: $L = 1(0)+2(0) = 0 = 0 \pmod 3$ 11: $L = 1(1)+2(1) = 3 = 0 \pmod 3$
01: $L = 1(0)+2(1) = 2 = 2 \pmod 3$ 21: $L = 1(2)+2(1) = 4 = 1 \pmod 3$
02: $L = 1(0)+2(2) = 4 = 1 \pmod 3$ 12: $L = 1(1)+2(2) = 5 = 2 \pmod 3$
10: $L = 1(1)+2(0) = 1 = 1 \pmod 3$ 22: $L = 1(2)+2(2) = 6 = 0 \pmod 3$
20: $L = 1(2)+2(0) = 2 = 2 \pmod 3$

En la figura 9-6 se muestran los bloques.

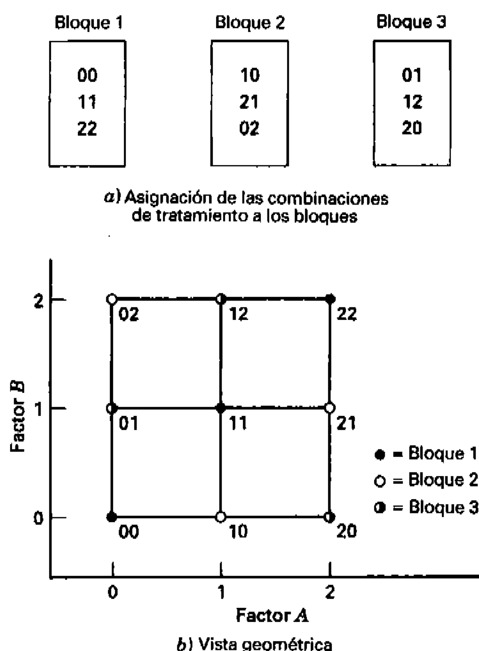


Figura 9-6 El diseño 3^2 en tres bloques con AB^2 confundida.

Los elementos del bloque principal forman un grupo con respecto a la adición módulo 3. Con referencia a la figura 9-6, se observa que 11 + 11 = 22 y 11 + 22 = 00. Las combinaciones de tratamientos de los otros dos bloques pueden generarse sumando, en módulo 3, cualquier elemento del nuevo bloque con los elementos del bloque principal. Por lo tanto, para el bloque 2 se usa 10 para obtener

$$10+00=10 \quad 10+11=21 \quad \text{y} \quad 10+22=02$$

Para generar el bloque 3, al utilizar 01, se encuentra

$$01+00=01 \quad 01+11=12 \quad \text{y} \quad 01+22=20$$

EJEMPLO 9-2 ··

El análisis estadístico del diseño 3^2 confundido en tres bloques se ilustra empleando los datos siguientes, los cuales provienen de la réplica única del diseño 3^2 que se muestra en la figura 9-6.

	Bloque 1	Bloque 2	Bloque 3
	00 = 4	10 = −2	01 = 5
	11 = −4	21 = 1	12 = −5
	22 = 0	02 = 8	20 = 0
Totales de los bloques =	0	7	0

Al aplicar los métodos convencionales para el análisis de factoriales, se encuentra que $SS_A = 131.56$ y $SS_B = 0.22$.

Tabla 9-4 Análisis de varianza de los datos del ejemplo 9-2

Fuente de variación	Suma de cuadrados	Grados de libertad
Bloques (AB^2)	10.89	2
A	131.56	2
B	0.22	2
AB	2.89	2
Total	145.56	8

Se encuentra también que

$$SS_{\text{Bloques}} = \frac{(0)^2 + (7)^2 + (0)^2}{3} - \frac{(7)^2}{9} = 10.89$$

Sin embargo, SS_{Bloques} es exactamente igual al componente AB^2 de la interacción. Para ver esto, las observaciones se escriben de la siguiente manera:

		Factor B		
		0	1	2
	0	4	5	8
Factor A	1	−2	−4	−5
	2	0	1	0

Recuerde, por la sección 9-1.2, que el componente I o AB^2 de la interacción AB puede encontrarse calculando la suma de cuadrados entre los totales de la diagonal de izquierda a derecha de la representación anterior. Se obtiene así

$$SS_{AB^2} = \frac{(0)^2 + (0)^2 + (7)^2}{3} - \frac{(7)^2}{9} = 10.89$$

valor que es idéntico a SS_{Bloques}.

El análisis de varianza se presenta en la tabla 9-4. Puesto que hay una sola réplica, no puede hacerse una prueba formal. No es una buena idea utilizar el componente AB de la interacción como una estimación del error.

....

Se considera ahora un diseño un poco más complicado; un diseño factorial 3^3 confundido en tres bloques con nueve corridas cada uno. El componente AB^2C^2 de la interacción de tres factores se confundirá con los bloques. La definición de contrastes es

$$L = x_1 + 2x_2 + 2x_3$$

Es sencillo verificar que las combinaciones de tratamientos 000, 012 y 101 se encuentran en el bloque principal. Las corridas restantes del bloque principal se generan de la siguiente manera:

(1) 000 (4) 101 + 101 = 202 (7) 101 + 021 = 122
(2) 012 (5) 012 + 012 = 021 (8) 012 + 202 = 211
(3) 101 (6) 101 + 012 = 110 (9) 021 + 202 = 220

Para encontrar las corridas de otro bloque se observa que la combinación de tratamientos 200 no está en el bloque principal. Por lo tanto, los elementos del bloque 2 son

(1) 200+000 = 200	(4) 200+202 = 102	(7) 200+122 = 022
(2) 200+012 = 212	(5) 200+021 = 221	(8) 200+211 = 111
(3) 200+101 = 001	(6) 200+110 = 010	(9) 200+220 = 120

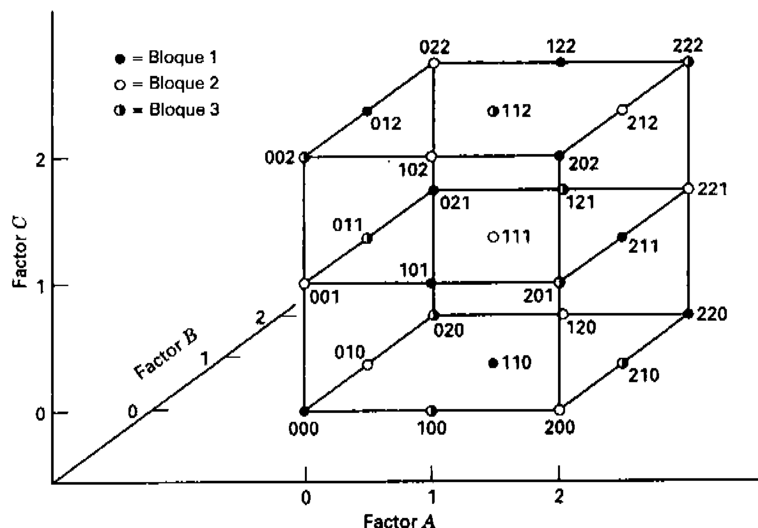
Observe que todas estas corridas satisfacen $L = 2 \pmod 3$. El último bloque se encuentra observando que 100 no pertenece al bloque 1 ni al bloque 2. Al usar 100 como arriba, se obtiene

(1) 100+000 = 100	(4) 100+202 = 002	(7) 100+122 = 222
(2) 100+012 = 112	(5) 100+021 = 121	(8) 100+211 = 011
(3) 100+101 = 201	(6) 100+110 = 210	(9) 100+220 = 020

Los bloques se ilustran en la figura 9-7.

Bloque 1	Bloque 2	Bloque 3
000	200	100
012	212	112
101	001	201
202	102	002
021	221	121
110	010	210
122	022	222
211	111	011
220	120	020

a) Asignación de las combinaciones de tratamientos a los bloques



b) Vista geométrica

Figura 9-7 El diseño 3^3 en tres bloques con AB^2C^2 confundida.

Tabla 9-5 Análisis de varianza de un diseño 3^3 con AB^2C^2 confundida

Fuente de variación	Grados de libertad
Bloques (AB^2C^2)	2
A	2
B	2
C	2
AB	4
AC	4
BC	4
Error ($ABC + AB^2C + ABC^2$)	6
Total	26

En la tabla 9-5 se presenta el análisis de varianza de este diseño. Al utilizar este esquema de confusión (o mezclado), se cuenta con información acerca de todos los efectos principales y las interacciones de dos factores. Los componentes restantes de la interacción de tres factores (ABC, AB^2C y ABC^2) se combinan como una estimación del error. La suma de cuadrados de esos tres componentes podría obtenerse por sustracción. En general, para el diseño 3^k en tres bloques se seleccionaría siempre un componente de la interacción de orden más alto para confundirlo con los bloques. Los demás componentes de esta interacción que no están confundidos pueden obtenerse calculando la interacción de k factores de la manera usual y restando de esta cantidad la suma de cuadrados de los bloques.

9-2.2 El diseño factorial 3^k en nueve bloques

En algunas situaciones experimentales puede ser necesario confundir el diseño 3^k en nueve bloques. Por lo tanto, ocho grados de libertad se confundirán con los bloques. Para construir estos diseños se eligen *dos* componentes de interacción y, como resultado, dos más se confundirán automáticamente, produciendo los ocho grados de libertad requeridos. Estos dos son las interacciones generalizadas de los dos efectos elegidos originalmente. En el sistema 3^k, las **interacciones generalizadas** de dos efectos (es decir, P y Q) se definen como PQ y PQ^2 (o P^2Q).

Los dos componentes de interacción elegidos inicialmente producen *dos* definiciones de contrastes

$$L_1 = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_k x_k = u \ (\text{mod } 3) \quad u = 0,\ 1,\ 2$$
$$L_2 = \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k = h \ (\text{mod } 3) \quad h = 0,\ 1,\ 2 \qquad (9\text{-}3)$$

donde $\{\alpha_i\}$ y $\{\beta_j\}$ son los exponentes de la primera y la segunda interacciones generalizadas, respectivamente, con la convención de que las primeras α_i y β_j diferentes de cero son la unidad. Las definiciones de contrastes de la ecuación 9-3 implican nueve ecuaciones simultáneas especificadas por el par de valores para L_1 y L_2. Las combinaciones de tratamientos que tienen el mismo par de valores para (L_1, L_2) se asignan al mismo bloque.

El bloque principal consta de las combinaciones de tratamientos que satisfacen $L_1 = L_2 = 0$ (mod 3). Los elementos de este bloque forman un grupo con respecto a la adición módulo 3; por lo tanto, el esquema presentado en la sección 9-2.1 puede usarse para generar los bloques.

Como un ejemplo, considere el diseño factorial 3^4 confundido en nueve bloques con nueve corridas cada uno. Suponga que se elige confundir ABC y AB^2D^2. Sus interacciones generalizadas

$$(ABC)(AB^2D^2) = A^2B^3CD^2 = (A^2B^3CD^2)^2 = AC^2D$$
$$(ABC)(AB^2D^2)^2 = A^3B^5CD^4 = B^2CD = (B^2CD)^2 = BC^2D^2$$

también están confundidas con los bloques. Las definiciones de contrastes de ABC y AB^2D^2 son

$$\begin{aligned} L_1 &= x_1 + x_2 + x_3 \\ L_2 &= x_1 + 2x_2 + 2x_4 \end{aligned} \qquad (9\text{-}4)$$

Los nueve bloques pueden construirse utilizando las definiciones de contrastes (ecuación 9-4) y la propiedad de la teoría de grupos del bloque principal. El diseño se muestra en la figura 9-8.

Para el diseño 3^k en nueve bloques habrá cuatro componentes de interacción confundidos. Los demás componentes de estas interacciones que no están confundidos pueden determinarse restando la suma de cuadrados del componente confundido de la suma de cuadrados de la interacción completa. El método descrito en la sección 9-1.3 puede ser útil para calcular los componentes de interacción.

9-2.3 El diseño factorial 3^k en 3^p bloques

El diseño factorial 3^k puede confundirse en 3^p bloques con 3^{k-p} observaciones cada uno, donde $p < k$. El procedimiento consiste en seleccionar p efectos independientes que habrán de confundirse con los bloques. Como resultado, exactamente otros $(3^p - 2p - 1)/2$ efectos se confunden de manera automática. Estos efectos son las interacciones generalizadas de los efectos elegidos originalmente.

Como una ilustración, considere un diseño 3^7 que va a confundirse en 27 bloques. Puesto que $p = 3$, se seleccionarían tres componentes de interacción independientes y se confundirían automáticamente otros $[3^3 - 2(3) - 1]/2 = 10$. Suponga que se eligen ABC^2DG, BCE^2F^2G y $BDEFG$. A partir de estos efectos pue-

Bloque 1	Bloque 2	Bloque 3	Bloque 4	Bloque 5	Bloque 6	Bloque 7	Bloque 8	Bloque 9
0000	0001	2000	0200	0020	0010	1000	0100	0002
0122	0120	2122	0022	0112	0102	1122	0222	0121
0211	0212	2211	0111	0201	0221	1211	0011	0210
1021	1022	0021	0221	1011	1001	2021	1121	1020
1110	1111	0110	1010	1100	1120	2110	1210	1112
1202	1200	0202	1102	1222	1212	2202	1002	1201
2012	2010	1012	2212	2002	2022	0012	2112	2011
2101	2012	1101	2001	2121	2111	0101	2201	2100
2220	2221	1220	2120	2210	2200	0220	2020	2222
$(L_1, L_2) = (0,0)$	$(0,1)$	$(2,2)$	$(2,0)$	$(2,1)$	$(1,2)$	$(1,1)$	$(1,0)$	$(0,2)$

Figura 9-8 El diseño 3^4 en nueve bloques con ABC, AB^2D^2, AC^2D y BC^2D^2 confundidas.

den construirse tres definiciones de contrastes, y los 27 bloques pueden generarse con los métodos descritos anteriormente. Los otros 10 efectos confundidos con los bloques son

$$(ABC^2DG)(BCE^2F^2G) = AB^2DE^2F^2G^2$$
$$(ABC^2DG)(BCE^2F^2G)^2 = AB^3C^4DE^4F^4G^3 = ACDEF$$
$$(ABC^2DG)(BDEFG) = AB^2C^2D^2EFG^2$$
$$(ABC^2DG)(BDEFG)^2 = AB^3C^2D^3E^2F^2G^3 = AC^2E^2F^2$$
$$(BCE^2F^2G)(BDEFG) = B^2CDE^3F^3G^2 = BC^2D^2G$$
$$(BCE^2F^2G)(BDEFG)^2 = B^3CD^2E^4F^4G^3 = CD^2EF$$
$$(ABC^2DG)(BCE^2F^2G)(BDEFG) = AB^3C^3D^2E^3F^3G^3 = AD^2$$
$$(ABC^2DG)^2(BCE^2F^2G)(BDEFG) = A^2B^4C^5D^3G^4 = AB^2CG^2$$
$$(ABC^2DG)(BCE^2F^2G)^2(BDEFG) = ABCD^2E^2F^2G$$
$$(ABC^2DG)(BCE^2F^2G)(BDEFG)^2 = ABC^3D^3E^4F^4G^4 = ABEFG$$

Se trata de un diseño enorme que requiere $3^7 = 2187$ observaciones dispuestas en 27 bloques con 81 observaciones cada uno.

9-3 RÉPLICAS FRACCIONADAS DEL DISEÑO FACTORIAL 3^k

El concepto de réplica fraccionada puede extenderse a los diseños factoriales 3^k. Debido a que una réplica completa del diseño 3^k puede requerir un número bastante grande de corridas incluso para valores moderados de k, las réplicas fraccionadas de estos diseños son de interés. Sin embargo, como se verá más adelante, algunos de estos diseños tienen estructuras de alias complicadas.

9-3.1 La fracción un tercio del diseño factorial 3^k

La fracción más grande del diseño 3^k es la fracción un tercio que contiene 3^{k-1} corridas. Por consiguiente, se hace referencia a él como el diseño factorial fraccionado 3^{k-1}. Para construir un diseño factorial fraccionado 3^{k-1} se selecciona un componente de interacción con dos grados de libertad (generalmente, la interacción de orden más alto) y se hace la partición del diseño 3^k *completo* en tres bloques. Cada uno de los tres bloques resultantes es un diseño fraccionado 3^{k-1} y puede seleccionarse cualquiera de los bloques para usarlo. Si $AB^{\alpha_2}C^{\alpha_3}\cdots K^{\alpha_k}$ es el componente de interacción utilizado para definir los bloques, entonces a $I = AB^{\alpha_2}C^{\alpha_3}\cdots K^{\alpha_k}$ se le llama la **relación de definición** del diseño factorial fraccionado. Cada efecto principal o componente de interacción estimado a partir del diseño 3^{k-1} tiene dos alias, los cuales pueden encontrarse multiplicando el efecto *tanto* por I *como* por I^2 módulo 3.

Como un ejemplo, considere una fracción un tercio del diseño 3^3. Puede seleccionarse cualquiera de los componentes de la interacción ABC para construir el diseño, es decir, ABC, AB^2C, ABC^2 o AB^2C^2. Por lo tanto, hay en realidad 12 fracciones un tercio *diferentes* del diseño 3^3 definidas por

$$x_1 + \alpha_2 x_2 + \alpha_3 x_3 = u \text{ (mod 3)}$$

donde $\alpha = 1$ o 2 y $u = 0, 1$ o 2. Suponga que se selecciona el componente de AB^2C^2. Cada fracción del diseño 3^{3-1} resultante contendrá exactamente $3^2 = 9$ combinaciones de tratamientos que deben satisfacer

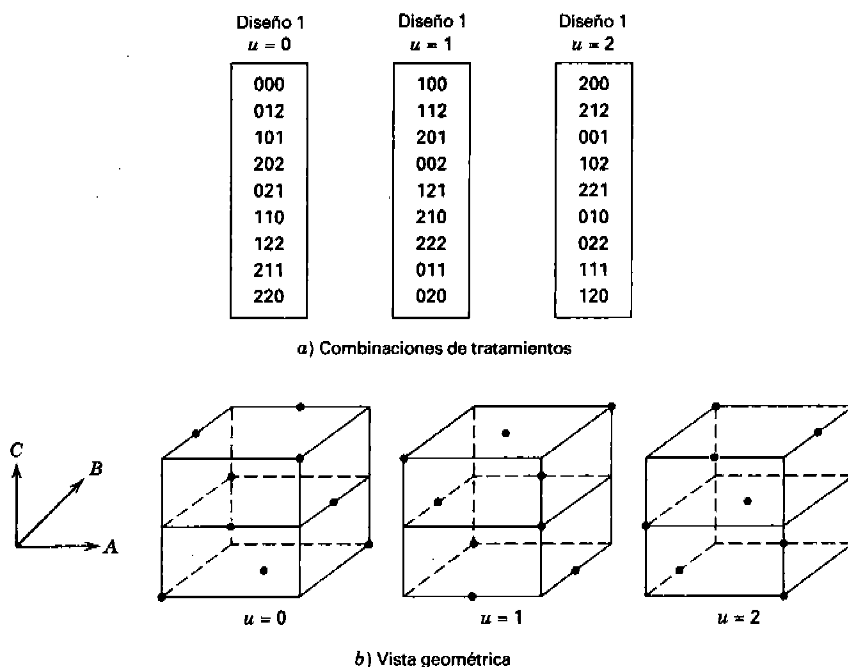
$$x_1 + 2x_2 + 2x_3 = u \text{ (mod 3)}$$

donde $u = 0, 1$ o 2. Es sencillo verificar que las tres fracciones un tercio son las que se muestran en la figura 9-9.

Si se corre cualquiera de los diseños 3^{3-1} de la figura 9-9, la estructura de los alias resultante es

$$A = A(AB^2C^2) = A^2B^2C^2 = ABC$$
$$A = A(AB^2C^2)^2 = A^3B^4C^4 = BC$$
$$B = B(AB^2C^2) = AB^3C^2 = AC^2$$
$$B = B(AB^2C^2)^2 = A^2B^5C^4 = ABC^2$$
$$C = C(AB^2C^2) = AB^2C^3 = AB^2$$
$$C = C(AB^2C^2)^2 = A^2B^4C^5 = AB^2C$$
$$AB = AB(AB^2C^2) = A^2B^3C^2 = AC$$
$$AB = AB(AB^2C^2)^2 = A^3B^5C^4 = BC^2$$

Por consiguiente, los cuatro efectos que en realidad se estiman a partir de los ocho grados de libertad del diseño son $A + BC + ABC$, $B + AC^2 + ABC^2$, $C + AB^2 + AB^2C$ y $AB + AC + BC^2$. Este diseño sólo tendría valor práctico si todas las interacciones fueran pequeñas en comparación con los efectos principales. Puesto que los efectos principales son alias de las interacciones de dos factores, se trata de un diseño de resolución III. Observe lo complejas que son las relaciones de los alias en este diseño. Cada efecto principal es alias de un *componente* de interacción. Si, por ejemplo, la interacción de dos factores BC es grande, esto distorsionará potencialmente la estimación del efecto principal de A y hará que sea muy complicada la interpretación del efecto de $AB + AC + BC^2$. Es muy difícil ver cómo este diseño podría ser de utilidad, a menos que se suponga que todas las interacciones son insignificantes.



Diseño 1 $u = 0$	Diseño 1 $u = 1$	Diseño 1 $u = 2$
000	100	200
012	112	212
101	201	001
202	002	102
021	121	221
110	210	010
122	222	022
211	011	111
220	020	120

a) Combinaciones de tratamientos

b) Vista geométrica

Figura 9-9 Las tres fracciones un tercio del diseño 3^3 con la relación de definición $I = AB^2C^2$.

Antes de dejar el diseño 3_{III}^{3-1}, observe que para el diseño con $u = 0$ (ver la figura 9-9), si se hace que A denote el renglón y B la columna, entonces el diseño puede escribirse como

000	012	021
101	110	122
202	211	220

que es un cuadrado latino 3 × 3. El supuesto de las interacciones insignificantes requerido para la interpretación única del diseño 3_{III}^{3-1} tiene su paralelo en el diseño del cuadrado latino. Sin embargo, los dos diseños surgen por motivos diferentes, uno como consecuencia de la réplica fraccionada y el otro de las restricciones sobre la aleatorización. Por la tabla 4-13 se observa que sólo hay 3 × 3 cuadrados latinos y que cada uno corresponde a uno de los doce diferentes diseños factoriales fraccionados 3^{3-1}.

Las combinaciones de tratamientos en un diseño 3^{k-1} con la relación de definición $I = AB^{\alpha_2}C^{\alpha_3}\cdots K^{\alpha_k}$ pueden construirse utilizando un método similar al que se empleó en la serie 2^{k-p}. Primero se escriben las 3^{k-1} corridas para un diseño factorial de tres niveles **completo** en $k - 1$ factores, con la notación común 0, 1, 2. Éste es el **diseño básico** en la terminología del capítulo 8. Después se introduce el factor k-ésimo igualando sus x_k niveles con el componente apropiado de la interacción de orden más alto, por ejemplo $AB^{\alpha_2}C^{\alpha_3}\cdots(K-1)^{\alpha_{k-1}}$, mediante la relación

$$x_k = \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_{k-1}x_{k-1} \tag{9-5}$$

donde $\beta_i = (3-\alpha_k)\alpha_i \pmod 3$ para $1 \le i \le k-1$. Se obtiene así un diseño con la resolución más alta posible.

Como una ilustración, se usa este método para generar el diseño 3_{IV}^{4-1} con la relación de definición $I = AB^2CD$ que se muestra en la tabla 9-6. Es sencillo verificar que los tres primeros dígitos de cada combinación de tratamientos de esta tabla son las 27 corridas de un diseño 3^3 completo. Se trata del diseño básico. Para AB^2CD se tiene $\alpha_1 = \alpha_3 = \alpha_4 = 1$ y $\alpha_2 = 2$. Esto implica que $\beta_1 = (3-1)\alpha_1 \pmod 3 = (3-1)(1) = 2$, $\beta_2 = (3-1)\alpha_2 \pmod 3 = (3-1)(2) = 4 = 1 \pmod 3$ y $\beta_3 = (3-1)\alpha_3 \pmod 3 = (3-1)(1) = 2$. Por lo tanto, la ecuación 9-5 queda como

$$x_4 = 2x_1 + x_2 + 2x_3 \tag{9-6}$$

Los niveles del cuarto factor satisfacen la ecuación 9-6. Por ejemplo, se tiene 2(0) + 1(0) + 2(0) = 0, 2(0) + 1(1) + 2(0) = 1, 2(1) + 1(1) + 2(0) = 3 = 0, etcétera.

El diseño 3_{IV}^{4-1} resultante tiene 26 grados de libertad que pueden usarse para calcular las sumas de cuadrados de los 13 efectos principales y los componentes de las interacciones (y sus alias). Los alias

Tabla 9-6 Un diseño 3_{IV}^{4-1} con $I = AB^2CD$

0000	0012	2221
0101	0110	0021
1100	0211	0122
1002	1011	0220
0202	1112	1020
1201	1210	1121
2001	2010	1222
2102	2111	2022
2200	2212	2120

de cualquier efecto se encuentran de la manera usual; por ejemplo, los alias de A son $A(AB^2CD) = ABC^2D^2$ y $A(AB^2CD)^2 = BC^2D^2$. Puede verificarse que los cuatro efectos principales están separados de cualquier componente de interacciones de dos factores, pero que algunos componentes de interacciones de dos factores son alias entre sí. Una vez más se observa la complejidad de la estructura de los alias. Si cualquiera de las interacciones de dos factores es grande, probablemente será muy difícil aislarla con este diseño.

El análisis estadístico de un diseño 3^{k-1} se lleva a cabo con los procedimientos usuales del análisis de varianza para experimentos factoriales. Las sumas de cuadrados de los componentes de interacciones pueden calcularse como en la sección 9-1. Cuando interprete los resultados, recuerde que los componentes de las interacciones no tienen ninguna interpretación práctica.

9-3.2 Otros diseños factoriales fraccionados 3^{k-p}

Para moderar los valores grandes de k, es deseable un fraccionamiento todavía mayor del diseño 3^k. En general, puede construirse una fracción $(\frac{1}{3})^p$ del diseño 3^k para $p < k$, donde la fracción contiene 3^{k-p} corridas. A este diseño se le llama el diseño factorial fraccionado 3^{k-p}. Por lo tanto, un diseño 3^{k-2} es una fracción un noveno, un diseño 3^{k-3} es una fracción un veintisieteavo, etcétera.

El procedimiento para construir un diseño factorial fraccionado 3^{k-p} consiste en seleccionar p componentes de interacciones y usar estos efectos para hacer la partición de las 3^k combinaciones de tratamientos en 3^p bloques. Entonces cada bloque es un diseño factorial fraccionado 3^{k-p}. La relación de definición I de cualquier fracción consta de los p efectos elegidos inicialmente y sus $(3^p - 2p - 1)/2$ interacciones generalizadas. El alias de cualquier efecto principal o componente de interacción se obtiene con la multiplicación módulo 3 del efecto por I e I^2.

Las corridas que definen un diseño factorial fraccionado 3^{k-p} también pueden generarse anotando primero las combinaciones de tratamientos de un diseño factorial 3^{k-p} completo e introduciendo después los p factores adicionales igualándolos con los componentes de las interacciones, como se hizo en la sección 9-3.1.

El procedimiento se ilustrará construyendo un diseño 3^{4-2}, es decir, una fracción un noveno del diseño 3^4. Sean AB^2C y BCD los dos componentes de interacciones elegidos para construir el diseño. Sus interacciones generalizadas son $(AB^2C)(BCD) = AC^2D$ y $(AB^2C)(BCD)^2 = ABD^2$. Por lo tanto, la relación de definición de este diseño es $I = AB^2C = BCD = AC^2D = ABD^2$, y el diseño es de resolución III. Las nueve combinaciones de tratamientos del diseño se encuentran apuntando un diseño 3^2 en los factores A y B, y agregando después dos nuevos factores haciendo

$$x_3 = 2x_1 + x_2$$
$$x_4 = 2x_2 + 2x_3$$

Tabla 9-7 Un diseño 3^{4-2}_{III} con $I = AB^2C$ e $I = BCD$

0000	0111	0222
1021	1102	1210
2012	2120	2201

Tabla 9-8 Estructura de los alias del diseño 3_{III}^{4-2} de la tabla 9-7

Efecto	Alias							
	I				I^2			
A	ABC^2	$ABCD$	ACD^2	AB^2D	BC^2	$AB^2C^2D^2$	CD^2	BD^2
B	AC	BC^2D^2	ABC^2D	AB^2D^2	ABC	CD	AB^2C^2D	AD^2
C	AB^2C^2	BC^2D	AD	$ABCD^2$	AB^2	BD	ACD	ABC^2D^2
D	AB^2CD	BCD^2	AC^2D^2	AB	AB^2CD^2	BC	AC^2	ABD

Esto es equivalente a usar AB^2C y BCD para hacer la partición del diseño 3^4 completo en nueve bloques y luego seleccionar uno de estos bloques como la fracción deseada. El diseño completo se muestra en la tabla 9-7.

Este diseño tiene ocho grados de libertad que pueden usarse para estimar cuatro efectos principales y sus alias. Los alias de cualquier efecto pueden encontrarse multiplicando el efecto módulo 3 por AB^2C, BCD, AC^2D, ABD^2 y sus cuadrados. En la tabla 9-8 se presenta la estructura de los alias completa del diseño.

Por la estructura de los alias se observa que este diseño sólo es útil en ausencia de interacciones. Además, si A denota los renglones y B las columnas, entonces al examinar la tabla 9-7 se observa que el diseño 3_{III}^{4-2} también es un cuadrado grecolatino.

El escrito de Connor y Zelen [28] contiene una extensa selección de diseños para $4 \leq k \leq 10$. Este escrito se elaboró para la National Bureau of Standards y es la tabla más completa disponible de los planes 3^{k-p}.

En esta sección se ha hecho notar en varias ocasiones la complejidad de las relaciones de los alias de los diseños factoriales fraccionados 3^{k-p}. En general, si k es moderadamente grande, por ejemplo $k \geq 4$ o 5, el tamaño del diseño 3^k llevará a muchos experimentadores a considerar fracciones bastante pequeñas. Desafortunadamente, estos diseños tienen relaciones de alias que incluyen **alias parciales** de componentes de interacciones con dos grados de libertad. Esto, a su vez, resulta en un diseño cuya interpretación será difícil, si no imposible, si las interacciones no son insignificantes. Además, no hay esquemas de aumento simples (como el doblez) que puedan usarse para combinar dos o más fracciones a fin de aislar las interacciones significativas. El uso del diseño 3^k suele sugerirse cuando hay curvatura presente. Sin embargo, hay alternativas más eficientes (ver el capítulo 11). Por estas razones, se puede concluir que los diseños factoriales fraccionados 3^{k-p} son soluciones que causan problemas; no son, en general, buenos diseños.

9-4 DISEÑOS FACTORIALES CON NIVELES MIXTOS

Se han resaltado los diseños factoriales y factoriales fraccionados en los que todos los factores tienen el mismo número de niveles. El sistema con dos niveles revisado en los capítulos 6, 7 y 8 es de particular utilidad. El sistema de tres niveles presentado en este capítulo es de utilidad mucho menor debido a que los diseños son relativamente grandes incluso para un número modesto de factores, y la mayoría de las fracciones pequeñas tienen relaciones de alias complejas que requerirían supuestos muy restrictivos respecto de las interacciones para ser útiles.

Estamos convencidos de que los diseños factoriales y factoriales fraccionados de dos niveles deberán ser la piedra angular de la experimentación industrial para el desarrollo de productos y procesos, detección de defectos y mejoramiento. Sin embargo, existen situaciones en las que es necesario incluir un factor (o algunos factores) que tiene más de dos niveles. Esto suele ocurrir cuando hay factores tanto cuantitativos como cualitativos en el experimento, y el factor cualitativo tiene (por ejemplo) tres niveles.

Tabla 9-9 Uso de factores con dos niveles para formar un factor con tres niveles

Factores con dos niveles		Factores con tres niveles
B	C	X
−	−	x_1
+	−	x_2
−	+	x_2
+	+	x_3

Si todos los factores son cuantitativos, entonces deberán usarse diseños de dos niveles con puntos centrales. En esta sección se indica cómo pueden incorporarse factores con tres y cuatro niveles en un diseño 2^k.

9-4.1 Factores con dos y tres niveles

Los diseños en los que algunos factores tienen dos niveles y otros tres niveles pueden derivarse de la tabla de signos positivos y negativos del diseño 2^k usual. El procedimiento general se ilustra mejor con un ejemplo. Suponga que se tienen dos variables, donde A tiene dos niveles y X tres. Considere la tabla de signos positivos y negativos del diseño 2^3 usual con ocho corridas. Los signos de las columnas B y C tienen el patrón que se muestra en el lado izquierdo de la tabla 9-9. Sea que los niveles de X estén representados por x_1, x_2 y x_3. En el lado derecho de la tabla 9-9 se muestra cómo se combinan los patrones de los signos de B y C para formar los niveles del factor con tres niveles.

Entonces el factor X tiene dos grados de libertad, y si el factor es cuantitativo, es posible hacer su partición en un componente lineal y uno cuadrático, con cada componente teniendo un grado de libertad. En la tabla 9-10 se muestra un diseño 2^3 con las columnas rotuladas para indicar los efectos reales que estiman, donde X_L y X_Q denotan los efectos lineal y cuadrático, respectivamente, de X. Observe que el efecto lineal de X es la suma de las estimaciones de los dos efectos calculadas a partir de las columnas asociadas generalmente con B y C, y que el efecto de A sólo puede calcularse a partir de las corridas donde X está en el nivel bajo o bien en el alto, es decir, las corridas 1, 2, 7 y 8. De manera similar, el efecto $A \times X_L$ es la suma de los dos efectos que se habrían calculado a partir de las columnas rotuladas generalmente AB y

Tabla 9-10 Un factor con dos niveles y un factor con tres niveles en un diseño 2^3

	A	X_L	X_L	$A \times X_L$	$A \times X_L$	X_Q	$A \times X_Q$	Combinaciones de tratamientos reales	
Corrida	A	B	C	AB	AC	BC	ABC	A	X
1	−	−	−	+	+	+	−	Bajo	Bajo
2	+	−	−	−	−	+	+	Alto	Bajo
3	−	+	−	−	+	−	+	Bajo	Intermedio
4	+	+	−	+	−	−	−	Alto	Intermedio
5	−	−	+	+	−	−	+	Bajo	Intermedio
6	+	−	+	−	+	−	−	Alto	Intermedio
7	−	+	+	−	−	+	−	Bajo	Alto
8	+	+	+	+	+	+	+	Alto	Alto

Tabla 9-11 Análisis de varianza del diseño de la tabla 9-10

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio
A	SS_A	1	MS_A
$X\ (X_L + X_Q)$	SS_X	2	MS_B
$AX\ (A \times X_L + A \times X_Q)$	SS_{AX}	2	MS_{AX}
Error (a partir de las corridas 3 y 5 y de las corridas 4 y 6)	SS_E	2	MS_E
Total	SS_T	7	

AC. Además, observe que las corridas 3 y 5 son réplicas. Por lo tanto, puede hacerse una estimación del error con un grado de libertad del error utilizando estas dos corridas. De manera similar, las corridas 4 y 6 son réplicas, y esto llevaría a una segunda estimación del error con un grado de libertad. La varianza promedio de estos dos pares de corridas podría usarse como cuadrado medio del error con dos grados de libertad. En la tabla 9-11 se resume el análisis de varianza completo.

Si se está dispuesto a suponer que las interacciones de dos factores y de órdenes superiores son insignificantes, el diseño de la tabla 9-10 puede convertirse en una fracción de resolución III con hasta cuatro factores con dos niveles y un solo factor con tres niveles. Esto se conseguiría asociando los factores de dos niveles con las columnas A, AB, AC y ABC. La columna BC no puede usarse para un factor de dos niveles porque contiene el efecto cuadrático del factor X de tres niveles.

Puede aplicarse el mismo procedimiento en los diseños 2^k de 16, 32 y 64 corridas. Para 16 corridas es posible construir factoriales fraccionados de resolución V con dos factores de dos niveles y con dos o tres factores de tres niveles. También puede obtenerse una fracción con 16 corridas de resolución V con 3 factores de dos niveles y un factor de tres niveles. Si se incluyen cuatro factores de dos niveles y un solo factor de tres niveles en 16 corridas, el diseño será de resolución III. Los diseños de 32 y 64 corridas permiten arreglos similares. Para un estudio adicional de algunos de estos diseños, ver Addleman [1b].

9-4.2 Factores con dos y cuatro niveles

Es muy sencillo incorporar un factor con cuatro niveles en un diseño 2^k. El procedimiento para hacerlo implica el uso de dos factores con dos niveles para representar el factor con cuatro niveles. Por ejemplo, suponga que A es un factor de cuatro niveles con los niveles a_1, a_2, a_3 y a_4. Considere dos columnas de la tabla usual de signos positivos y negativos, por ejemplo las columnas P y Q. El patrón de los signos de estas dos columnas se muestra en el lado izquierdo de la tabla 9-12. El lado derecho de esta tabla muestra cómo estos cuatro patrones de signos corresponderían con los cuatro niveles del factor A. Los efectos represen-

Tabla 9-12 El factor A con cuatro niveles expresado como dos factores con dos niveles

	Factores con dos niveles		Factores con cuatro niveles
Corrida	P	Q	A
1	−	−	a_1
2	+	−	a_2
3	−	+	a_3
4	+	+	a_4

Tabla 9-13 Un factor con cuatro niveles y 2 factores con dos niveles en 16 corridas

Corrida	(*A*	*B*)	= *X*	*C*	*D*	*AB*	*AC*	*BC*	*ABC*	*AD*	*BD*	*ABD*	*CD*	*ACD*	*BCD*	*ABCD*
1	−	−	x_1	−	−	+	+	+	−	+	+	−	+	−	−	+
2	+	−	x_2	−	−	−	−	+	+	−	+	+	+	+	−	−
3	−	+	x_3	−	−	−	+	−	+	+	−	+	+	−	+	−
4	+	+	x_4	−	−	+	−	−	−	−	−	−	+	+	+	+
5	−	−	x_1	+	−	+	−	−	+	+	+	−	−	+	+	−
6	+	−	x_2	+	−	−	+	−	−	−	+	+	−	−	+	+
7	−	+	x_3	+	−	−	−	+	−	+	−	+	−	+	−	+
8	+	+	x_4	+	−	+	+	+	+	−	−	−	−	−	−	−
9	−	−	x_1	−	+	+	+	+	−	−	−	+	−	+	+	−
10	+	−	x_2	−	+	−	−	+	+	+	−	−	−	−	+	+
11	−	+	x_3	−	+	−	+	−	+	−	+	−	−	+	−	+
12	+	+	x_4	−	+	+	−	−	−	+	+	+	−	−	−	−
13	−	−	x_1	+	+	+	−	−	+	−	−	+	+	−	−	+
14	+	−	x_2	+	+	−	+	−	−	+	−	−	+	+	−	−
15	−	+	x_3	+	+	−	−	+	−	−	+	−	+	−	+	−
16	+	+	x_4	+	+	+	+	+	+	+	+	+	+	+	+	+

tados por las columnas P y Q y la interacción PQ son mutuamente ortogonales y corresponden al efecto de A con tres grados de libertad.

Para ilustrar esta idea con mayor detalle, suponga que se tiene un factor de cuatro niveles y dos factores de dos niveles y que es necesario estimar todos los efectos principales y las interacciones en las que intervienen estos factores. Esto puede hacerse con un diseño de 16 corridas. En la tabla 9-13 se presenta la tabla usual de signos positivos y negativos del diseño 2^4 con 16 corridas, donde las columnas A y B se usan para formar el factor de cuatro niveles, por ejemplo X, con los niveles x_1, x_2, x_3 y x_4. Se calcularían las sumas de cuadrados de cada columna $A, B, ..., ABCD$ exactamente igual que en el sistema 2^k usual. Después las sumas de cuadrados de todos los factores X, C, D y sus interacciones se forman de la manera siguiente:

$$SS_X = SS_A + SS_B + SS_{AB} \quad \text{(3 grados de libertad)}$$
$$SS_C = SS_C \quad \text{(1 grado de libertad)}$$
$$SS_D = SS_D \quad \text{(1 grado de libertad)}$$
$$SS_{CD} = SS_{CD} \quad \text{(1 grado de libertad)}$$
$$SS_{XC} = SS_{AC} + SS_{BC} + SS_{ABC} \quad \text{(3 grados de libertad)}$$
$$SS_{XD} = SS_{AD} + SS_{BD} + SS_{ABD} \quad \text{(3 grados de libertad)}$$
$$SS_{XCD} = SS_{ACD} + SS_{BCD} + SS_{ABCD} \quad \text{(3 grados de libertad)}$$

A este diseño podría llamársele 4×2^2. Si uno está dispuesto a ignorar las interacciones de dos factores, pueden asociarse hasta nueve factores adicionales de dos niveles con la columna de la interacción de dos factores (excepto AB), la columna de la interacción de tres factores y la columna de la interacción de cuatro factores.

9-5 PROBLEMAS

9-1. Se estudian los efectos de la fuerza del revelador (A) y el tiempo de revelado (B) sobre la densidad de la película de placa fotográfica. Se usan tres fuerzas y tres tiempos, y se corren cuatro réplicas de un experimento factorial 3^2. Los datos de este experimento se presentan a continuación. Analizar los datos utilizando los métodos estándares para experimentos factoriales.

	Tiempo de revelado (minutos)					
Fuerza del revelador	10		14		18	
1	0	2	1	3	2	5
	5	4	4	2	4	6
2	4	6	6	8	9	10
	7	5	7	7	8	5
3	7	10	10	10	12	10
	8	7	8	7	9	8

9-2. Calcular los componentes I y J de la interacción de dos factores del problema 9-1.

9-3. Se llevó a cabo un experimento para estudiar el efecto de tres tipos diferentes de botellas de 32 onzas (A) y tres tipos diferentes de aparadores de venta (B) —anaqueles permanentes lisos, aparadores al final del pasillo con anaqueles enrejados y refrigeradores para refrescos— sobre el tiempo que toma acomodar diez cajas de 12 botellas en los aparadores. Se usaron tres empleados (factor C) en el experimento, y se corrieron dos réplicas de un diseño factorial 3^3. Los datos del tiempo observado se muestran en la tabla siguiente. Analizar los datos y sacar conclusiones.

		Réplica I			Réplica II		
Empleado	Tipo de botella	Permanente	Final del pasillo	Refrigerador	Permanente	Final del pasillo	Refrigerador
1	Plástico	3.45	4.14	5.80	3.36	4.19	5.23
	Vidrio de 28 mm	4.07	4.38	5.48	3.52	4.26	4.85
	Vidrio de 38 mm	4.20	4.26	5.67	3.68	4.37	5.58
2	Plástico	4.80	5.22	6.21	4.40	4.70	5.88
	Vidrio de 28 mm	4.52	5.15	6.25	4.44	4.65	6.20
	Vidrio de 38 mm	4.96	5.17	6.03	4.39	4.75	6.38
3	Plástico	4.08	3.94	5.14	3.65	4.08	4.49
	Vidrio de 28 mm	4.30	4.53	4.99	4.04	4.08	4.59
	Vidrio de 38 mm	4.17	4.86	4.85	3.88	4.48	4.90

9-4. Un investigador médico estudia el efecto de la lidocaína sobre el nivel de enzimas en el músculo cardiaco de perros beagle. En el experimento se usan tres marcas comerciales de lidocaína (A), tres dosis (B) y tres perros (C), y se corren dos réplicas de un diseño factorial 3^3. Los niveles de enzimas observados se presentan a continuación. Analizar los datos de este experimento.

		Réplica I			Réplica II		
		Perro			Perro		
Marca de lidocaína	Fuerza de la dosis	1	2	3	1	2	3
1	1	96	84	85	84	85	86
	2	94	99	98	95	97	90
	3	101	106	98	105	104	103
2	1	85	84	86	80	82	84
	2	95	98	97	93	99	95
	3	108	114	109	110	102	100
3	1	84	83	81	83	80	79
	2	95	97	93	92	96	93
	3	105	100	106	102	111	108

9-5. Calcular los componentes I y J de las interacciones de dos factores del ejemplo 10-1.

9-6. Se realiza un experimento en un proceso químico utilizando un diseño factorial 3^2. Los factores del diseño son la temperatura y la presión, y la variable de respuesta es el rendimiento. Los datos que resultan de este experimento se presentan a continuación:

	Presión, psig		
Temperatura, °C	100	120	140
80	47.58, 48.77	64.97, 69.22	80.92, 72.60
90	51.86, 82.43	88.47, 84.23	93.95, 88.54
100	71.18, 92.77	96.57, 88.72	76.58, 83.04

a) Analizar los datos de este experimento conduciendo un análisis de varianza. ¿Qué conclusiones pueden sacarse?

b) Analizar gráficamente los residuales. ¿Hay algún motivo de preocupación respecto de los supuestos subyacentes o de la adecuación del modelo?

c) Verificar que si se hace que los niveles bajo, intermedio y alto de ambos factores de este diseño asuman los niveles −1, 0 y +1, entonces un ajuste de mínimos cuadrados de un modelo de segundo orden del rendimiento es

$$\hat{y} = 86.81 + 10.4x_1 + 8.42x_2 - 7.17x_1^2 - 7.86x_2^2 - 7.69x_1x_2$$

d) Confirmar que el modelo del inciso *c* puede escribirse en términos de las variables naturales —la temperatura (T) y la presión (P)— como

$$\hat{y} = -1335.63 + 18.56T + 8.59P - 0.072T^2 - 0.0196P^2 - 0.0384TP$$

e) Construir una gráfica de contorno del rendimiento como una función de la presión y la temperatura. Con base en el examen de esta gráfica, ¿dónde se recomendaría operar este proceso?

9-7. *a*) Confundir un diseño 3^3 en tres bloques utilizando el componente ABC^2 de la interacción de tres factores. Comparar los resultados obtenidos con el diseño de la figura 9-7.

b) Confundir un diseño 3^3 en tres bloques utilizando el componente AB^2C de la interacción de tres factores. Comparar los resultados con el diseño de la figura 9-7.

c) Confundir un diseño 3^3 en tres bloques utilizando el componente ABC de la interacción de tres factores. Comparar los resultados obtenidos con el diseño de la figura 9-7.

d) Después de observar los diseños de los incisos *a*, *b* y *c* y la figura 9-7, ¿qué conclusiones pueden sacarse?

9-8. Confundir un diseño 3^4 en tres bloques utilizando el componente AB^2CD de la interacción de cuatro factores.

9-9. Considere los datos de la primera réplica del problema 9-3. Suponiendo que no fue posible hacer las 27 observaciones el mismo día, establecer un diseño para conducir el experimento en tres días con AB^2C confundida con los bloques. Analizar los datos.

9-10. Delinear la tabla del análisis de varianza del diseño 3^4 en nueve bloques. ¿Se trata de un diseño práctico?

9-11. Considere los datos del problema 9-3. Si ABC está confundida en la réplica I y ABC^2 está confundida en la réplica II, realizar el análisis de varianza.

9-12. Considere los datos de la réplica I del problema 9-3. Suponga que sólo se corre una fracción un tercio de este diseño con $I = ABC$. Construir el diseño, determinar la estructura de los alias y analizar los datos.

9-13. Por el examen de la figura 9-9, ¿qué tipo de diseño quedaría si después de completar las nueve primeras corridas pudiera eliminarse uno de los tres factores?

9-14. Construir un diseño 3_{IV}^{4-1} con $I = ABCD$. Escribir la estructura de los alias de este diseño.

9-15. Verificar que el diseño del problema 9-14 es un diseño de resolución IV.

9-16. Construir un diseño 3^{5-2} con $I = ABC$ e $I = CDE$. Escribir la estructura de los alias de este diseño. ¿Cuál es la resolución de este diseño?

9-17. Construir un diseño 3^{9-6} y verificar que es un diseño de resolución III.

9-18. Construir un diseño 4×2^3 confundido en dos bloques con 16 observaciones cada uno. Delinear el análisis de varianza de este diseño.

9-19. Delinear la tabla del análisis de varianza de un diseño factorial $2^2 3^2$. Comentar la manera en que este diseño puede confundirse en bloques.

9-20. Empezando con un diseño 2^4 de 16 corridas, indicar cómo pueden incorporarse dos factores de tres niveles en este experimento. ¿Cuántos factores de dos niveles pueden incluirse si se quiere cierta información sobre las interacciones de dos factores?

9-21. Empezando con un diseño 2^4 de 16 corridas, indicar cómo pueden incorporarse un factor con tres niveles y tres factores con dos niveles, de tal modo que siga siendo posible la estimación de las interacciones de dos factores.

9-22. En el problema 8-26 el lector conoció a Harry y Judy Peterson-Nedry, dos amigos del autor que son propietarios de un viñedo y una fábrica vinícola en Newberg, Oregon. En ese problema se describió la aplicación de diseños factoriales fraccionados de dos niveles en su producto Pinot Noir 1985. En 1987 quisieron conducir otro experimento Pinot Noir. Las variables de este experimento fueron

Variable	Niveles
Clon de Pinot Noir	Wadenswil, Pommard
Tamaño de la uva	Pequeño, grande
Temperatura de fermentación	80°F, 85°F, 90/80°F, 90°F
Uva completa	Ninguno, 10%
Tiempo de maceración	10 días, 21 días
Tipo de levadura	Assmanhau, Champagne
Tipo de roble	Tronçais, Allier

Harry y Judy decidieron usar un diseño factorial fraccionado de dos niveles con 16 corridas, tratando los cuatro niveles de la temperatura de fermentación como dos variables de dos niveles. Como en el problema 8-26, utilizaron las calificaciones de un panel de catadores como variable de respuesta. El diseño y las calificaciones promedio resultantes se presentan enseguida:

Corrida	Clon	Tamaño de la uva	Temperatura de fermentación		Uva completa	Tiempo de maceración	Tipo de levadura	Tipo de roble	Calificación promedio
1	−	−	−	−	−	−	−	−	4
2	+	−	−	−	−	+	+	+	10
3	−	+	−	−	+	−	+	+	6
4	+	+	−	−	+	+	−	−	9
5	−	−	+	−	+	+	+	−	11
6	+	−	+	−	+	−	−	+	1
7	−	+	+	−	−	+	−	+	15
8	+	+	+	−	−	−	+	−	5
9	−	−	−	+	+	+	−	+	12
10	+	−	−	+	+	−	+	−	2
11	−	+	−	+	−	+	+	−	16
12	+	+	−	+	−	−	−	+	3
13	−	−	+	+	−	−	+	+	8
14	+	−	+	+	−	+	−	−	14
15	−	+	+	+	+	−	−	−	7
16	+	+	+	+	+	+	+	+	13

a) Describir los alias de este diseño.
b) Analizar los datos y sacar conclusiones.
c) ¿Qué comparaciones pueden hacerse entre este experimento y el experimento del Pinot Noir 1985 del problema 8-26?

9-23. En un artículo de W.D. Baten publicado en el volumen de 1956 de *Industrial Quality Control* se describe un experimento para estudiar el efecto de tres factores sobre la longitud de unas barras de acero. Cada barra se sometió a uno de dos procesos de tratamiento térmico y se cortó en una de cuatro máquinas en una de tres horas durante el día (8 a.m., 11 a.m. o 3 p.m.). Los datos de la longitud codificada son los siguientes:

Hora del día	Proceso de tratamiento térmico	Máquina							
		1		2		3		4	
8 a.m.	1	6	9	7	9	1	2	6	6
		1	3	5	5	0	4	7	3
	2	4	6	6	5	−1	0	4	5
		0	1	3	4	0	1	5	4

Hora del día	Proceso de tratamiento térmico	Máquina 1		Máquina 2		Máquina 3		Máquina 4	
11 a.m.	1	6	3	8	7	3	2	7	9
		1	–1	4	8	1	0	11	6
	2	3	1	6	4	2	0	9	4
		1	–2	1	3	–1	1	6	3
3 p.m.	1	5	4	10	11	–1	2	10	5
		9	6	6	4	6	1	4	8
	2	6	0	8	7	0	–2	4	3
		3	7	10	0	4	–4	7	0

a) Analizar los datos de este experimento suponiendo que las cuatro observaciones de cada celda son réplicas.

b) Analizar los residuales de este experimento. ¿Existe algún indicio de que hay un punto atípico en una de las celdas? Si se encuentra un punto atípico, eliminarlo y repetir el análisis del inciso *a*. ¿A qué conclusiones se llega?

c) Suponga que las observaciones de las celdas son las longitudes (codificadas) de barras que se procesaron conjuntamente en el tratamiento térmico y después se cortaron secuencialmente (es decir, en orden) en las cuatro máquinas. Analizar los datos y determinar los efectos de los tres factores sobre la longitud promedio.

d) Calcular la varianza logarítmica de las observaciones de cada celda. Analizar esta respuesta. ¿Qué conclusiones pueden sacarse?

e) Suponga que la hora en que se corta una barra en realidad no puede controlarse durante la producción rutinaria. Analizar la longitud promedio y la varianza logarítmica de la longitud de cada una de las 12 barras cortadas en cada combinación máquina/proceso de tratamiento térmico. ¿Qué conclusiones pueden sacarse?

10 Ajuste de modelos de regresión

10-1 INTRODUCCIÓN

En muchos problemas hay dos o más variables relacionadas, y el interés se centra en modelar y explorar esta relación. Por ejemplo, en un proceso químico el rendimiento del producto está relacionado con la temperatura de operación. Quizá el ingeniero químico quiera construir un modelo que relacione el rendimiento con la temperatura para usarlo después como herramienta de predicción o bien de optimización o control del proceso.

En general, suponga que hay una sola **variable dependiente** o **de respuesta** y que depende de k **variables independientes** o **regresores,** por ejemplo, $x_1, x_2, ..., x_k$. La relación que existe entre estas variables se caracteriza por un modelo matemático llamado **modelo de regresión.** Dicho modelo se ajusta a un conjunto de datos muestrales. En ocasiones el experimentador conoce la forma exacta de la verdadera relación funcional entre y y $x_1, x_2, ..., x_k$, por ejemplo $y = \phi(x_1, x_2, ..., x_k)$. Sin embargo, en la mayoría de los casos no se conoce la verdadera relación funcional, y el experimentador elige una función apropiada para aproximar ϕ. Los modelos polinomiales de orden inferior son de uso generalizado como funciones de aproximación.

Existe una fuerte relación recíproca entre el diseño de experimentos y el análisis de regresión. A lo largo de este libro se ha destacado la importancia de expresar cuantitativamente los resultados de un experimento, en términos del **modelo empírico,** a fin de facilitar su comprensión, interpretación e implementación. Los modelos de regresión constituyen la base para conseguirlo. Se ha presentado en múltiples ocasiones el modelo de regresión que representaba los resultados de un experimento. En este capítulo se presentan algunos aspectos del ajuste de estos modelos. Presentaciones más completas de la regresión se encuentran en Montgomery y Peck [82] y Myers [84].

Los métodos de regresión se utilizan con frecuencia para analizar datos de **experimentos no planeados,** como podría ser el caso de la observación de fenómenos no controlados o de registros históricos. Los métodos de regresión también son muy útiles en experimentos diseñados cuando algo "salió mal". En este capítulo se ilustran algunas de estas situaciones.

10-2 MODELOS DE REGRESIÓN LINEAL

La atención se centrará en el ajuste de modelos de regresión lineal. Para ilustrar, suponga que quiere desarrollarse un modelo empírico que relacione la viscosidad de un polímero con la temperatura y la velocidad de alimentación del catalizador. Un modelo que podría describir esta relación es

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon \tag{10-1}$$

donde y representa la viscosidad, x_1 la temperatura y x_2 la velocidad de alimentación del catalizador. Se trata de un **modelo de regresión lineal múltiple** con dos variables independientes. Es común llamar a las variables independientes **variables predictoras** o **regresores (variables de regresión).** Se utiliza el término *lineal* porque la ecuación 10-1 es una función lineal de los parámetros desconocidos β_0, β_1 y β_2. El modelo describe un plano en el espacio bidimensional x_1, x_2. El parámetro β_0 define la intersección del plano con el eje de las ordenadas. En ocasiones β_1 y β_2 se denominan los *coeficientes de regresión parcial*, porque β_1 mide el cambio esperado en y para cada cambio unitario de x_1 cuando x_2 se mantiene constante, y β_2 mide el cambio esperado en y para cada cambio unitario de x_2 cuando x_1 se mantiene constante.

En general, la variable de respuesta y puede relacionarse con k regresores. Al modelo

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \varepsilon \tag{10-2}$$

se le llama *modelo de regresión lineal múltiple* con k regresores. A los parámetros β_j, $j = 0, 1, \ldots, k$ se les llama los *coeficientes de regresión*. Este modelo describe un hiperplano en el espacio de k dimensiones de los regresores $\{x_j\}$. El parámetro β_j representa el cambio esperado en la respuesta y para un cambio unitario en x_j cuando las variables independientes restantes x_i $(i \neq j)$ se mantienen constantes.

Con frecuencia los modelos cuya apariencia es más compleja que la ecuación 10-2 pueden también analizarse mediante técnicas de regresión lineal múltiple. Por ejemplo, considere la incorporación de un término de interacción en el modelo de primer orden en dos variables, por ejemplo

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \varepsilon \tag{10-3}$$

Si se hace $x_3 = x_1 x_2$ y $\beta_3 = \beta_{12}$, entonces la ecuación 10-3 puede escribirse como

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon \tag{10-4}$$

que es un modelo de regresión lineal múltiple estándar con tres regresores. Recuerde que en algunos ejemplos de los capítulos 6, 7 y 8 se presentaron varios modelos empíricos similares a las ecuaciones 10-2 y 10-4 para expresar cuantitativamente los resultados de un diseño factorial de dos niveles. Como otro ejemplo, considere el **modelo de superficie de respuesta** de segundo orden en dos variables:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{12} x_1 x_2 + \varepsilon \tag{10-5}$$

Si se hace $x_3 = x_1^2, x_4 = x_2^2, x_5 = x_1 x_2, \beta_3 = \beta_{11}, \beta_4 = \beta_{22}$ y $\beta_5 = \beta_{12}$, entonces esta expresión queda como

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \varepsilon \tag{10-6}$$

que es un modelo de regresión lineal. Este modelo se ha visto también en ejemplos anteriores de este libro. En general, cualquier modelo de regresión que es lineal en los parámetros (los valores β) es un modelo de regresión lineal, independientemente de la forma de la superficie de respuesta que genera.

En este capítulo se resumirán los métodos para estimar los parámetros de los modelos de regresión lineal múltiple. A este procedimiento suele llamársele el **ajuste del modelo.** Se analizarán también los métodos para probar hipótesis y para construir intervalos de confianza para estos modelos, así como para

verificar la adecuación del ajuste del modelo. La atención se centra en los aspectos del análisis de regresión que son útiles en los experimentos diseñados. Para presentaciones más completas de la regresión, referirse a Montgomery y Peck [82] y Myers [84].

10-3 ESTIMACIÓN DE LOS PARÁMETROS EN MODELOS DE REGRESIÓN LINEAL

El método de mínimos cuadrados se usa de manera típica para estimar los coeficientes de regresión de un modelo de regresión lineal múltiple. Suponga que se cuenta con $n > k$ observaciones de la variable de respuesta, por ejemplo, $y_1, y_2, ..., y_n$. Junto con cada respuesta observada y_i se tendrá una observación de cada uno de los regresores, y sea que x_{ij} denote la observación o nivel i-ésimo de la variable x_j. Los datos aparecerán como en la tabla 10-1. Se supone que el término del error ε del modelo tiene $E(\varepsilon) = 0$ y $V(\varepsilon) = \sigma^2$ y que las $\{\varepsilon_i\}$ son variables aleatorias no correlacionadas.

La ecuación del modelo (ecuación 10-2) puede escribirse en términos de las observaciones de la tabla 10-1 como

$$\begin{aligned} y_i &= \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + \varepsilon_i \\ &= \beta_0 + \sum_{j=1}^{k} \beta_j x_{ij} + \varepsilon_i \qquad i = 1, 2, ..., n \end{aligned} \tag{10-7}$$

El método de mínimos cuadrados consiste en elegir las β de la ecuación 10-7 de tal modo que la suma de cuadrados de los errores, ε_i, se minimice. La función de mínimos cuadrados es

$$\begin{aligned} L &= \sum_{i=1}^{n} \varepsilon_i^2 \\ &= \sum_{i=1}^{n} \left(y_i - \beta_0 - \sum_{j=1}^{k} \beta_j x_{ij} \right)^2 \end{aligned} \tag{10-8}$$

La función L debe minimizarse con respecto a $\beta_0, \beta_1, ..., \beta_k$. Los estimadores de mínimos cuadrados, por ejemplo $\hat{\beta}_0, \hat{\beta}_1, ..., \hat{\beta}_k$, deben satisfacer

$$\left. \frac{\partial L}{\partial \beta_0} \right|_{\hat{\beta}_0, \hat{\beta}_1, ..., \hat{\beta}_k} = -2 \sum_{i=1}^{n} \left(y_i - \hat{\beta}_0 - \sum_{j=1}^{k} \hat{\beta}_j x_{ij} \right) = 0 \tag{10-9a}$$

y

$$\left. \frac{\partial L}{\partial \beta_j} \right|_{\hat{\beta}_0, \hat{\beta}_1, ..., \hat{\beta}_k} = -2 \sum_{i=1}^{n} \left(y_i - \hat{\beta}_0 - \sum_{j=1}^{k} \hat{\beta}_j x_{ij} \right) x_{ij} = 0 \qquad j = 1, 2, ..., k \tag{10-9b}$$

Tabla 10-1 Datos de una regresión lineal múltiple

y	x_1	x_2	...	x_k
y_1	x_{11}	x_{12}	...	x_{1k}
y_2	x_{21}	x_{22}	...	x_{2k}
⋮	⋮	⋮		⋮
y_n	x_{n1}	x_{n2}	...	x_{nk}

Al simplificar la ecuación 10-9, se obtiene

$$\begin{aligned}
n\hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^{n} x_{i1} \quad + \hat{\beta}_2 \sum_{i=1}^{n} x_{i2} \quad + \cdots + \hat{\beta}_k \sum_{i=1}^{n} x_{ik} \quad &= \sum_{i=1}^{n} y_i \\
\hat{\beta}_0 \sum_{i=1}^{n} x_{i1} + \hat{\beta}_1 \sum_{i=1}^{n} x_{i1}^2 \quad + \hat{\beta}_2 \sum_{i=1}^{n} x_{i1}x_{i2} + \cdots + \hat{\beta}_k \sum_{i=1}^{n} x_{i1}x_{ik} &= \sum_{i=1}^{n} x_{i1}y_i \\
\vdots \qquad \qquad \vdots \qquad \qquad \vdots \qquad \qquad \vdots \qquad &\quad \vdots \\
\hat{\beta}_0 \sum_{i=1}^{n} x_{ik} + \hat{\beta}_1 \sum_{i=1}^{n} x_{ik}x_{i1} + \hat{\beta}_2 \sum_{i=1}^{n} x_{ik}x_{i2} + \cdots + \hat{\beta}_k \sum_{i=1}^{n} x_{ik}^2 \quad &= \sum_{i=1}^{n} x_{ik}y_i
\end{aligned} \tag{10-10}$$

Estas ecuaciones se denominan **ecuaciones normales de mínimos cuadrados.** Observe que hay $p = k + 1$ ecuaciones normales, una para cada uno de los coeficientes de regresión desconocidos. La solución de las ecuaciones normales serán los estimadores de mínimos cuadrados de los coeficientes de regresión $\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_k$.

Es más sencillo resolver las ecuaciones normales si se expresan en la notación matricial. A continuación se presenta el desarrollo matricial de las ecuaciones normales que es análogo al desarrollo de la ecuación 10-10. El modelo en términos de las observaciones, ecuación 10-7, puede escribirse en notación matricial como

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

donde

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \qquad \mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \ldots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix},$$

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} \quad \text{y} \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

En general, **y** es un vector $(n \times 1)$ de las observaciones, **X** es una matriz $(n \times p)$ de los niveles de las variables independientes, $\boldsymbol{\beta}$ es un vector $(p \times 1)$ de los coeficientes de regresión, y $\boldsymbol{\varepsilon}$ es un vector $(n \times 1)$ de los errores aleatorios.

Quiere encontrarse el vector de los estimadores de mínimos cuadrados, $\hat{\boldsymbol{\beta}}$, que minimice

$$L = \sum_{i=1}^{n} \varepsilon_i^2 = \boldsymbol{\varepsilon}'\boldsymbol{\varepsilon} = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$$

Observe que L puede expresarse como

$$\begin{aligned} L &= \mathbf{y}'\mathbf{y} - \boldsymbol{\beta}'\mathbf{X}'\mathbf{y} - \mathbf{y}'\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta} \\ &= \mathbf{y}'\mathbf{y} - 2\boldsymbol{\beta}'\mathbf{X}'\mathbf{y} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta} \end{aligned} \tag{10-11}$$

ya que $\boldsymbol{\beta}'\mathbf{X}'\mathbf{y}$ es una matriz (1×1), o un escalar, y su transpuesta $(\boldsymbol{\beta}'\mathbf{X}'\mathbf{y})' = \mathbf{y}'\mathbf{X}\boldsymbol{\beta}$ es el mismo escalar. Los estimadores de mínimos cuadrados deben satisfacer

$$\left.\frac{\partial L}{\partial \boldsymbol{\beta}}\right|_{\hat{\beta}} = -2\mathbf{X}'\mathbf{y} + 2\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{0}$$

cuya simplificación es

$$\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{y} \tag{10-12}$$

La ecuación 10-12 es la forma matricial de las ecuaciones normales de mínimos cuadrados. Es idéntica a la ecuación 10-10. Para resolver las ecuaciones normales, ambos miembros de la ecuación 10-12 se multiplican por la inversa de $\mathbf{X}'\mathbf{X}$. Por lo tanto, el estimador de mínimos cuadrados de $\boldsymbol{\beta}$ es

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \tag{10-13}$$

Es sencillo ver que la forma matricial de las ecuaciones normales es idéntica a la forma escalar. Al desarrollar en detalle la ecuación 10-12, se obtiene

$$\begin{bmatrix} n & \sum_{i=1}^{n} x_{i1} & \sum_{i=1}^{n} x_{i2} & \cdots & \sum_{i=1}^{n} x_{ik} \\ \sum_{i=1}^{n} x_{i1} & \sum_{i=1}^{n} x_{i1}^2 & \sum_{i=1}^{n} x_{i1}x_{i2} & \cdots & \sum_{i=1}^{n} x_{i1}x_{ik} \\ \vdots & \vdots & \vdots & & \vdots \\ \sum_{i=1}^{n} x_{ik} & \sum_{i=1}^{n} x_{ik}x_{i1} & \sum_{i=1}^{n} x_{ik}x_{i2} & \cdots & \sum_{i=1}^{n} x_{ik}^2 \end{bmatrix} \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_k \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_{i1}y_i \\ \vdots \\ \sum_{i=1}^{n} x_{ik}y_i \end{bmatrix}$$

Si se efectúa la multiplicación matricial indicada, se obtendrá la forma escalar de las ecuaciones normales (es decir, la ecuación 10-10). En esta forma es sencillo ver que $\mathbf{X}'\mathbf{X}$ es una matriz simétrica $(p \times p)$ y que $\mathbf{X}'\mathbf{y}$ es un vector columna $(p \times 1)$. Observe la estructura especial de la matriz $\mathbf{X}'\mathbf{X}$. Los elementos de la diagonal de $\mathbf{X}'\mathbf{X}$ son las sumas de cuadrados de los elementos de las columnas de $\mathbf{X}$, y los elementos que no están en la diagonal son las sumas de los productos cruzados de los elementos de las columnas de $\mathbf{X}$. Además, observe que los elementos de $\mathbf{X}'\mathbf{y}$ son las sumas de los productos cruzados de las columnas de $\mathbf{X}$ y las observaciones $\{y_i\}$.

El modelo de regresión ajustado es

$$\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}} \tag{10-14}$$

En notación escalar, el modelo ajustado es

$$\hat{y}_i = \hat{\beta}_0 + \sum_{i=1}^{k} \hat{\beta}_j x_{ij} \qquad i = 1, 2, \ldots, n$$

La diferencia entre la observación real y_i y el valor ajustado correspondiente $\hat{y}_i$ es el **residual,** es decir, $e_i = y_i - \hat{y}_i$. El vector $(n \times 1)$ de los residuales se denota por

$$\mathbf{e} = \mathbf{y} - \hat{\mathbf{y}} \tag{10-15}$$

Estimación de σ^2

Por lo general también es necesario estimar σ^2. Para desarrollar un estimador de este parámetro, considere la suma de cuadrados de los residuales, por ejemplo

$$\begin{aligned} SS_E &= \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \\ &= \sum_{i=1}^{n} e_i^2 \\ &= \mathbf{e}'\mathbf{e} \end{aligned}$$

Al sustituir $\mathbf{e} = \mathbf{y} - \hat{\mathbf{y}} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}$, se tiene

$$\begin{aligned} SS_E &= (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}) \\ &= \mathbf{y}'\mathbf{y} - \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} - \mathbf{y}'\mathbf{X}\hat{\boldsymbol{\beta}} + \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} \\ &= \mathbf{y}'\mathbf{y} - 2\hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} + \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} \end{aligned}$$

Puesto que $\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{y}$, esta última ecuación queda como

$$SS_E = \mathbf{y}'\mathbf{y} - \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} \tag{10-16}$$

A la ecuación 10-16 se le llama la **suma de cuadrados residual** o **del error,** y tiene $n - p$ grados de libertad asociados con ella. Puede demostrarse que

$$E(SS_E) = \sigma^2(n - p)$$

por lo que un estimador insesgado de σ^2 está dado por

$$\hat{\sigma}^2 = \frac{SS_E}{n - p} \tag{10-17}$$

Propiedades de los estimadores

El método de mínimos cuadrados produce un estimador insesgado del parámetro $\boldsymbol{\beta}$ del modelo de regresión lineal. Esto puede demostrarse fácilmente tomando el valor esperado de $\hat{\boldsymbol{\beta}}$ de la siguiente manera:

$$\begin{aligned} E(\hat{\boldsymbol{\beta}}) &= E[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}] \\ &= E[(\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon})] \\ &= E[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}\boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}] \\ &= \boldsymbol{\beta} \end{aligned}$$

ya que $E(\boldsymbol{\varepsilon}) = \mathbf{0}$ y $(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X} = \mathbf{I}$. Por lo tanto, $\hat{\boldsymbol{\beta}}$ es un estimador insesgado de $\boldsymbol{\beta}$.

La propiedad de la varianza de $\hat{\boldsymbol{\beta}}$ se expresa en la **matriz de covarianza:**

$$\text{Cov}(\hat{\boldsymbol{\beta}}) \equiv E\left\{[\hat{\boldsymbol{\beta}} - E(\hat{\boldsymbol{\beta}})][\hat{\boldsymbol{\beta}} - E(\hat{\boldsymbol{\beta}})]'\right\} \tag{10-18}$$

que es una matriz simétrica cuyo elemento i-ésimo de la diagonal principal es la varianza del coeficiente de regresión individual $\hat{\beta}_i$, y cuyo elemento (ij)-ésimo es la covarianza entre $\hat{\beta}_i$ y $\hat{\beta}_j$. La matriz de covarianza de $\hat{\boldsymbol{\beta}}$ es

$$\text{Cov}(\hat{\boldsymbol{\beta}}) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1} \tag{10-19}$$

EJEMPLO 10-1 ....

En la tabla 10-2 se muestran 16 observaciones de la viscosidad de un polímero (y) y dos variables del proceso: la temperatura de reacción (x_1) y la velocidad de alimentación del catalizador (x_2). Se ajustará el modelo de regresión lineal múltiple

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$$

Tabla 10-2 Datos de la viscosidad del ejemplo 10-1 (viscosidad en centistokes @ 100°C)

Observación	Temperatura (x_1, °C)	Velocidad de alimentación del catalizador (x_2, lb/h)	Viscosidad
1	80	8	2256
2	93	9	2340
3	100	10	2426
4	82	12	2293
5	90	11	2330
6	99	8	2368
7	81	8	2250
8	96	10	2409
9	94	12	2364
10	93	11	2379
11	97	13	2440
12	95	11	2364
13	100	8	2404
14	85	12	2317
15	86	9	2309
16	87	12	2328

a estos datos. La matriz **X** y el vector **y** son

$$\mathbf{X} = \begin{bmatrix} 1 & 80 & 8 \\ 1 & 93 & 9 \\ 1 & 100 & 10 \\ 1 & 82 & 12 \\ 1 & 90 & 11 \\ 1 & 99 & 8 \\ 1 & 81 & 8 \\ 1 & 96 & 10 \\ 1 & 94 & 12 \\ 1 & 93 & 11 \\ 1 & 97 & 13 \\ 1 & 95 & 11 \\ 1 & 100 & 8 \\ 1 & 85 & 12 \\ 1 & 86 & 9 \\ 1 & 87 & 12 \end{bmatrix} \qquad \mathbf{y} = \begin{bmatrix} 2256 \\ 2340 \\ 2426 \\ 2293 \\ 2330 \\ 2368 \\ 2250 \\ 2409 \\ 2364 \\ 2379 \\ 2440 \\ 2364 \\ 2404 \\ 2317 \\ 2309 \\ 2328 \end{bmatrix}$$

La matriz **X′X** es

$$\mathbf{X'X} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 80 & 93 & \cdots & 87 \\ 8 & 9 & \cdots & 12 \end{bmatrix} \begin{bmatrix} 1 & 80 & 8 \\ 1 & 93 & 9 \\ \vdots & \vdots & \vdots \\ 1 & 87 & 12 \end{bmatrix}$$

$$= \begin{bmatrix} 16 & 1458 & 164 \\ 1458 & 133{,}560 & 14{,}946 \\ 164 & 14{,}946 & 1{,}726 \end{bmatrix}$$

y el vector $\mathbf{X}'\mathbf{y}$ es

$$\mathbf{X}'\mathbf{y} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 80 & 93 & \cdots & 87 \\ 8 & 9 & \cdots & 12 \end{bmatrix} \begin{bmatrix} 2256 \\ 2340 \\ \vdots \\ 2328 \end{bmatrix}$$

$$= \begin{bmatrix} 37{,}577 \\ 3{,}429{,}550 \\ 385{,}562 \end{bmatrix}$$

La estimación de mínimos cuadrados de $\boldsymbol{\beta}$ es

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$$

o

$$\hat{\boldsymbol{\beta}} = \begin{bmatrix} 14.176004 & -0.129746 & -0.223453 \\ -0.129746 & 1.429184 \times 10^{-3} & -4.763947 \times 10^{-5} \\ -0.223453 & -4.763947 \times 10^{-5} & 2.222381 \times 10^{-2} \end{bmatrix} \begin{bmatrix} 37{,}577 \\ 3{,}429{,}550 \\ 385{,}562 \end{bmatrix}$$

$$= \begin{bmatrix} 1566.07777 \\ 7.62129 \\ 8.58485 \end{bmatrix}$$

El ajuste de mínimos cuadrados, con los coeficientes de regresión expresados con dos cifras decimales, es

Tabla 10-3 Valores predichos, residuales y otros diagnósticos del ejemplo 10-1

Observación i	y_i	Valor predicho $\hat{y}_i$	Residual e_i	h_{ii}	Residual studentizado	D_i	R-student
1	2256	2244.5	11.5	0.350	0.87	0.137	0.87
2	2340	2352.1	−12.1	0.102	−0.78	0.023	−0.77
3	2426	2414.1	11.9	0.177	0.80	0.046	0.79
4	2293	2294.0	−1.0	0.251	−0.07	0.001	−0.07
5	2330	2346.4	−16.4	0.077	−1.05	0.030	−1.05
6	2368	2389.3	−21.3	0.265	−1.52	0.277	−1.61
7	2250	2252.1	−2.1	0.319	−0.15	0.004	−0.15
8	2409	2383.6	25.4	0.098	1.64	0.097	1.76
9	2364	2385.5	−21.5	0.142	−1.42	0.111	−1.48
10	2379	2369.3	9.7	0.080	0.62	0.011	0.60
11	2440	2416.9	23.1	0.278	1.66	0.354	1.80
12	2364	2384.5	−20.5	0.096	−1.32	0.062	−1.36
13	2404	2396.9	7.1	0.289	0.52	0.036	0.50
14	2317	2316.9	0.1	0.185	0.01	0.000	<0.01
15	2309	2298.8	10.2	0.134	0.67	0.023	0.66
16	2328	2332.1	−4.1	0.156	−0.28	0.005	−0.27

Tabla 10-4 Salida de Minitab para el modelo de regresión de la viscosidad, ejemplo 10-1

Análisis de regresión

```
The regression equation is
Viscosity = 1566 + 7.62 Temp + 8.58 Feed Rate

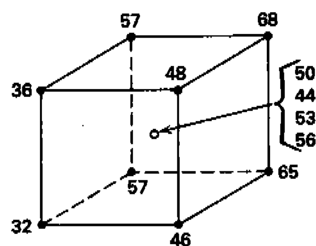
Predictor          Coef       StDev          T          P
Constant        1566.08       61.59      25.43      0.000
Temp             7.6213      0.6184      12.32      0.000
Feed Rat          8.585       2.439       3.52      0.004

S = 16.36        R-Sq = 92.7%     R-Sq(adj) = 91.6%

Analysis of Variance

Source              DF          SS          MS          F          P
Regression           2       44157       22079      82.50      0.000
Residual Error      13        3479         268
Total               15       47636

Source         DF      Seq SS
Temp            1       40841
Feed Rat        1        3316
```



	Variables del proceso			Variables codificadas			Rendimiento,
Corrida	Temperatura (°C)	Presión (psig)	Concentración (g/l)	x_1	x_2	x_3	y
1	120	40	15	−1	−1	−1	32
2	160	40	15	1	−1	−1	46
3	120	80	15	−1	1	−1	57
4	160	80	15	1	1	−1	65
5	120	40	30	−1	−1	1	36
6	160	40	30	1	−1	1	48
7	120	80	30	−1	1	1	57
8	160	80	30	1	1	1	68
9	140	60	22.5	0	0	0	50
10	140	60	22.5	0	0	0	44
11	140	60	22.5	0	0	0	53
12	140	60	22.5	0	0	0	56

$$x_1 = \frac{\text{Temperatura} - 140}{20}, \quad x_2 = \frac{\text{Presión} - 60}{20}, \quad x_3 = \frac{\text{Concentración} - 22.5}{7.5}$$

Figura 10-5 Diseño experimental del ejemplo 10-2.

Ajuste de modelos de regresión en experimentos diseñados

Se ha usado con frecuencia un modelo de regresión para presentar los resultados de un experimento diseñado en una forma cuantitativa. Se ofrece ahora un ejemplo completo donde se indica cómo se hace esto. Se presentan enseguida otros tres ejemplos breves que ilustran otras aplicaciones útiles del análisis de regresión en los experimentos diseñados.

EJEMPLO 10-2 ....

Análisis de regresión de un diseño factorial 2^3

Un ingeniero químico está investigando el rendimiento de un proceso. Tres de las variables del proceso son de interés: la temperatura, la presión y la concentración del catalizador. Cada variable puede correrse en un nivel bajo y uno alto, y el ingeniero decide correr un diseño 2^3 con cuatro puntos centrales. En la figura 10-5 se muestra el diseño y los rendimientos resultantes, donde se presentan tanto los niveles naturales del diseño como la notación de variables codificadas $+1, -1$ que se utiliza normalmente en los diseños factoriales 2^k para representar los niveles de los factores.

Suponga que el ingeniero decide ajustar un modelo que sólo incluye los efectos principales, por ejemplo

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon$$

Para este modelo, la matriz **X** y el vector **y** son

$$\mathbf{X} = \begin{bmatrix} 1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \qquad \mathbf{y} = \begin{bmatrix} 32 \\ 46 \\ 57 \\ 65 \\ 36 \\ 48 \\ 57 \\ 68 \\ 50 \\ 44 \\ 53 \\ 56 \end{bmatrix}$$

Es sencillo demostrar que

$$\mathbf{X'X} = \begin{bmatrix} 12 & 0 & 0 & 0 \\ 0 & 8 & 0 & 0 \\ 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 8 \end{bmatrix} \qquad \mathbf{X'y} = \begin{bmatrix} 612 \\ 45 \\ 85 \\ 9 \end{bmatrix}$$

Puesto que **X'X** es *diagonal*, el inverso que se requiere también es diagonal, y las estimaciones de mínimos cuadrados de los coeficientes de regresión son

$$\hat{\boldsymbol{\beta}} = (\mathbf{X'X})^{-1}\mathbf{X'y} = \begin{bmatrix} 1/12 & 0 & 0 & 0 \\ 0 & 1/8 & 0 & 0 \\ 0 & 0 & 1/8 & 0 \\ 0 & 0 & 0 & 1/8 \end{bmatrix} \begin{bmatrix} 612 \\ 45 \\ 85 \\ 9 \end{bmatrix} = \begin{bmatrix} 51.000 \\ 5.625 \\ 10.625 \\ 1.125 \end{bmatrix}$$

El modelo de regresión ajustado es

$$\hat{y} = 51.000 + 5.625x_1 + 10.625x_2 + 1.125x_3$$

Como se ha hecho uso de ellos en muchas ocasiones, los coeficientes de regresión guardan una estrecha relación con las estimaciones de los efectos que se obtendrían por el análisis usual de un diseño 2^3. Por ejemplo, el efecto de la temperatura es (referirse a la figura 10-5)

$$\begin{aligned} T &= \bar{y}_{T^+} - \bar{y}_{T^-} \\ &= 56.75 - 45.50 \\ &= 11.25 \end{aligned}$$

Observe que el coeficiente de regresión de x_1 es

$$(11.25)/2 = 5.625$$

Es decir, el coeficiente de regresión es exactamente la mitad de la estimación usual del efecto. Esto siempre se cumplirá para un diseño 2^k. Como se señaló antes, en los capítulos 6 al 8 se empleó este resultado para producir modelos de regresión, valores ajustados y residuales en varios experimentos de dos niveles. Este ejemplo demuestra que las estimaciones de los efectos de un diseño 2^k son estimaciones de mínimos cuadrados.

En el ejemplo 10-2 es sencillo obtener la matriz inversa porque $\mathbf{X}'\mathbf{X}$ es diagonal. Intuitivamente, esto parece ofrecer ventajas, no sólo porque los cálculos se simplifican sino también porque los estimadores de todos los coeficientes de regresión no están correlacionados, es decir, $\text{Cov}(\hat{\beta}_i, \hat{\beta}_j) = 0$. Si los niveles de las variables x pueden elegirse antes de recabar los datos, quizá sea deseable diseñar el experimento de tal modo que resulte una $\mathbf{X}'\mathbf{X}$ diagonal.

En la práctica puede ser relativamente sencillo conseguir esto. Se sabe que los elementos de $\mathbf{X}'\mathbf{X}$ que están fuera de la diagonal son las sumas de los productos cruzados de las columnas en $\mathbf{X}$. Por lo tanto, es necesario hacer que el producto interior de las columnas de $\mathbf{X}$ sean iguales a cero; es decir, estas columnas deben ser **ortogonales.** A los diseños experimentales que poseen esta propiedad para ajustar un modelo de regresión se les llama **diseños ortogonales.** En general, el diseño factorial 2^k es un diseño ortogonal para ajustar el modelo de regresión lineal múltiple.

Los métodos de regresión son en extremo útiles cuando algo "sale mal" en un experimento diseñado. Esto se ilustra en los dos ejemplos siguientes.

EJEMPLO 10-3

Un diseño factorial 2^3 con una observación faltante

Considere el diseño factorial 2^3 con cuatro puntos centrales del ejemplo 10-2. Suponga que cuando se realizó este experimento, faltó la corrida con todas las variables en el nivel alto (la corrida 8 de la figura 10-5). Esto puede ocurrir por varias razones: el sistema de medición puede producir una lectura incorrecta, la combinación de los niveles de los factores quizá no sea la apropiada, la unidad experimental puede estar dañada, etcétera.

Se ajustará el modelo de los efectos principales

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon$$

utilizando las 11 observaciones restantes. La matriz $\mathbf{X}$ y el vector $\mathbf{y}$ son

$$\mathbf{X}=\begin{bmatrix}1 & -1 & -1 & -1\\ 1 & 1 & -1 & -1\\ 1 & -1 & 1 & -1\\ 1 & 1 & 1 & -1\\ 1 & -1 & -1 & 1\\ 1 & 1 & -1 & 1\\ 1 & -1 & 1 & 1\\ 1 & 0 & 0 & 0\\ 1 & 0 & 0 & 0\\ 1 & 0 & 0 & 0\\ 1 & 0 & 0 & 0\end{bmatrix} \quad \mathbf{y}=\begin{bmatrix}32\\ 46\\ 57\\ 65\\ 36\\ 48\\ 57\\ 50\\ 44\\ 53\\ 56\end{bmatrix}$$

Para estimar los parámetros del modelo se forman

$$\mathbf{X}'\mathbf{X}=\begin{bmatrix}11 & -1 & -1 & -1\\ -1 & 7 & -1 & -1\\ -1 & -1 & 7 & -1\\ -1 & -1 & -1 & 7\end{bmatrix} \quad \mathbf{X}'\mathbf{y}=\begin{bmatrix}544\\ -23\\ 17\\ -59\end{bmatrix}$$

y entonces

$$\hat{\boldsymbol{\beta}}=(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$$

$$=\begin{bmatrix}9.61538\times10^{-2} & 1.92307\times10^{-2} & 1.92307\times10^{-2} & 1.92307\times10^{-2}\\ 1.92307\times10^{-2} & 0.15385 & 2.88462\times10^{-2} & 2.88462\times10^{-2}\\ 1.92307\times10^{-2} & 2.88462\times10^{-2} & 0.15385 & 2.88462\times10^{-2}\\ 1.92307\times10^{-2} & 2.88462\times10^{-2} & 2.88462\times10^{-2} & 0.15385\end{bmatrix}\begin{bmatrix}544\\ -23\\ 17\\ -59\end{bmatrix}$$

$$=\begin{bmatrix}51.25\\ 5.75\\ 10.75\\ 1.25\end{bmatrix}$$

Por lo tanto, el modelo ajustado es

$$\hat{y}=51.25+5.75x_1+10.75x_2+1.25x_3$$

Compare este modelo con el que se obtuvo en el ejemplo 10-2, donde se usaron las 12 observaciones. Los coeficientes de regresión son muy similares. Debido a la estrecha relación entre los coeficientes de regresión y los efectos de los factores, las conclusiones no sufrirían una alteración sustancial por la observación faltante. Sin embargo, observe que las estimaciones de los efectos han dejado de ser ortogonales, ya que $(\mathbf{X}'\mathbf{X})$ y su inversa ya no son diagonales.

EJEMPLO 10-4

Niveles imprecisos de los factores del diseño

Cuando se corre un experimento diseñado, en ocasiones es difícil alcanzar y mantener los niveles precisos de los factores requeridos por el diseño. Las discrepancias pequeñas no son importantes, pero las grandes son motivo de preocupación potencial. Los métodos de regresión son útiles en el análisis de un

Tabla 10-5 Diseño experimental del ejemplo 10-4

	Variables del proceso			Variables codificadas			
Corrida	Temperatura (°C)	Presión (psig)	Concentración (g/l)	x_1	x_2	x_3	Rendimiento y
1	125	41	14	–0.75	–0.95	–1.133	32
2	158	40	15	0.90	–1	–1	46
3	121	82	15	–0.95	1.1	–1	57
4	160	80	15	1	1	–1	65
5	118	39	33	–1.10	–1.05	1.14	36
6	163	40	30	1.15	–1	1	48
7	122	80	30	–0.90	1	1	57
8	165	83	30	1.25	1.15	1	68
9	140	60	22.5	0	0	0	50
10	140	60	22.5	0	0	0	44
11	140	60	22.5	0	0	0	53
12	140	60	22.5	0	0	0	56

experimento diseñado cuando el experimentador no ha podido obtener los niveles requeridos de los factores.

Para ilustrar, el experimento de la tabla 10-5 presenta una variación del diseño 2^3 del ejemplo 10-2, donde muchas de las combinaciones de prueba no son exactamente las que se especifican en el diseño. Las dificultades parecen haber ocurrido sobre todo con la variable temperatura.

Se ajustará el modelo de los efectos principales

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon$$

La matriz **X** y el vector **y** son

$$\mathbf{X} = \begin{bmatrix} 1 & -0.75 & -0.95 & -1.133 \\ 1 & 0.90 & -1 & -1 \\ 1 & -0.95 & 1.1 & -1 \\ 1 & 1 & 1 & -1 \\ 1 & -1.10 & -1.05 & 1.4 \\ 1 & 1.15 & -1 & 1 \\ 1 & -0.90 & 1 & 1 \\ 1 & 1.25 & 1.15 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \qquad \mathbf{y} = \begin{bmatrix} 32 \\ 46 \\ 57 \\ 65 \\ 36 \\ 48 \\ 57 \\ 68 \\ 50 \\ 44 \\ 53 \\ 56 \end{bmatrix}$$

Para estimar los parámetros del modelo se necesitan

$$\mathbf{X'X} = \begin{bmatrix} 12 & 0.60 & 0.25 & 0.2670 \\ 0.60 & 8.18 & 0.31 & -0.1403 \\ 0.25 & 0.31 & 8.5375 & -0.3437 \\ 0.2670 & -0.1403 & -0.3437 & 9.2437 \end{bmatrix} \qquad \mathbf{X'y} = \begin{bmatrix} 612 \\ 77.55 \\ 161.50 \\ 19.144 \end{bmatrix}$$

Entonces

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$$

$$= \begin{bmatrix} 8.37447\times10^{-2} & -6.09871\times10^{-3} & -2.33542\times10^{-3} & -2.59833\times10^{-3} \\ -6.09871\times10^{-3} & 0.12289 & -4.20766\times10^{-3} & 1.88490\times10^{-3} \\ -2.33542\times10^{-3} & -4.20766\times10^{-3} & 0.11753 & 4.37851\times10^{-3} \\ -2.59833\times10^{-3} & 1.88490\times10^{-3} & 4.37851\times10^{-3} & 0.10845 \end{bmatrix} \begin{bmatrix} 612 \\ 77.55 \\ 161.50 \\ 19.144 \end{bmatrix}$$

$$= \begin{bmatrix} 50.36496 \\ 5.41932 \\ 10.16672 \\ 1.07653 \end{bmatrix}$$

El modelo de regresión ajustado, con los coeficientes reportados con dos cifras decimales, es

$$\hat{y} = 50.36 + 2x_1 + 10.17x_2 + 1.08x_3$$

Al comparar este resultado con el modelo original del ejemplo 10-2, donde los niveles de los factores fueron exactamente los que se especificaron en el diseño, se observa muy poca diferencia. La interpretación práctica de los resultados de este experimento no sufriría alteraciones sustanciales por la incapacidad del experimentador para alcanzar exactamente los niveles deseados de los factores.

EJEMPLO 10-5

Separación de alias de interacciones en un diseño factorial fraccionado

En el capítulo 8 se señaló la posibilidad de separar los alias de las interacciones de un diseño factorial fraccionado mediante el proceso llamado doblez o plegado. Para un diseño de resolución III, un plegado completo se construye corriendo una segunda fracción en la que los signos están invertidos respecto de los signos de la fracción original. Entonces el diseño combinado puede usarse para separar los alias de todos los efectos principales de las interacciones de dos factores.

Una dificultad con el plegado es que requiere un segundo grupo de corridas de tamaño idéntico al del diseño original. Por lo general es posible separar los alias de ciertas interacciones de interés aumentando el diseño original con un número de corridas menor que las que se requieren en un plegado completo. Los métodos de regresión son una forma fácil de formular este problema y de ver cómo puede resolverse.

Para ilustrar, suponga que se ha corrido un diseño 2_{IV}^{4-1}. En la tabla 8-3 se muestra la fracción principal de este diseño, en la que $I = ABCD$. Suponga que después de que se observaron los datos de los ocho primeros ensayos, los efectos más grandes fueron A, B, C, D (se ignoran las interacciones de tres factores que son alias de estos efectos principales), y la cadena de alias $AB + CD$. Las otras dos cadenas de alias pueden ignorarse, pero es claro que AB o CD o ambas interacciones de dos factores son grandes. Para dilucidar cuáles son las interacciones importantes podría, desde luego, correrse la fracción alterna, para lo cual se requerirían otros ocho ensayos. Entonces las 16 corridas podrían usarse para estimar los efectos principales y las interacciones de dos factores.

Es posible separar los alias de AB y CD en un número de ensayos adicionales menor que ocho. Suponga que quiere ajustarse el modelo

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_{12} x_1 x_2 + \beta_{34} x_3 x_4 + \varepsilon$$

donde x_1, x_2, x_3 y x_4 son las variables codificadas que representan a A, B, C y D. Utilizando el diseño de la tabla 8-3, la matriz **X** de este modelo es

$$
\mathbf{X} = \begin{array}{c} \\ \end{array}\overset{\begin{array}{ccccccc} & x_1 & x_2 & x_3 & x_4 & x_1x_2 & x_3x_4 \end{array}}{\begin{bmatrix} 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & 1 & -1 & -1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 & 1 & -1 & -1 \\ 1 & 1 & 1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & 1 & -1 & -1 & -1 \\ 1 & -1 & 1 & 1 & -1 & -1 & -1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}}
$$

donde se han anotado las variables arriba de las columnas a fin de facilitar la comprensión. Observe que la columna x_1x_2 es idéntica a la columna x_3x_4 (como se anticipaba, ya que AB o x_1x_2 es alias de CD o x_3x_4), lo cual implica una dependencia lineal en las columnas de **X**. Por lo tanto, no pueden estimarse tanto β_{12} como β_{34} en el modelo. Sin embargo, suponga que se agrega la corrida única $x_1 = -1, x_2 = -1, x_3 = -1$ y $x_4 = 1$ de la fracción alterna a las ocho corridas originales. Entonces la matriz **X** del modelo queda como

$$
\mathbf{X} = \overset{\begin{array}{ccccccc} & x_1 & x_2 & x_3 & x_4 & x_1x_2 & x_3x_4 \end{array}}{\begin{bmatrix} 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & 1 & -1 & -1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 & 1 & -1 & -1 \\ 1 & 1 & 1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & 1 & -1 & -1 & -1 \\ 1 & -1 & 1 & 1 & -1 & -1 & -1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & -1 & 1 & 1 & -1 \end{bmatrix}}
$$

Observe que ahora las columnas x_1x_2 y x_3x_4 ya no son idénticas, y el modelo puede ajustarse incluyendo a las dos interacciones x_1x_2 (AB) y x_3x_4 (CD). Las magnitudes de los coeficientes de regresión brindarán información respecto a cuáles son las interacciones importantes.

Aun cuando al agregar una sola corrida se separarán los alias de las interacciones AB y CD, este enfoque tiene una desventaja. Suponga que existe un efecto de tiempo (o un efecto de bloque) entre las ocho primeras corridas y la última corrida que se agregó arriba. Al agregarse una columna a la matriz **X** para los bloques, se obtiene lo siguiente:

$$
\mathbf{X} = \overset{\begin{array}{cccccccc} & x_1 & x_2 & x_3 & x_4 & x_1x_2 & x_3x_4 & \text{bloques} \end{array}}{\begin{bmatrix} 1 & -1 & -1 & -1 & -1 & 1 & 1 & -1 \\ 1 & 1 & -1 & -1 & 1 & -1 & -1 & -1 \\ 1 & -1 & 1 & -1 & 1 & -1 & -1 & -1 \\ 1 & 1 & 1 & -1 & -1 & 1 & 1 & -1 \\ 1 & -1 & -1 & 1 & 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & -1 \\ 1 & -1 & -1 & -1 & 1 & 1 & -1 & 1 \end{bmatrix}}
$$

Se ha supuesto que el factor del bloque estaba en el nivel bajo o "–" durante las ocho primeras corridas, y en el nivel alto o "+" durante la novena corrida. Es sencillo ver que la suma de los productos cruzados de cada columna con la columna del bloque no es cero, lo cual significa que los bloques han dejado de ser ortogonales para los tratamientos, o que el efecto del bloque afecta ahora a las estimaciones de los coeficientes de regresión del modelo. Para conseguir la ortogonalidad de los bloques, debe agregarse un número par de corridas. Por ejemplo, con las cuatro corridas

x_1	x_2	x_3	x_4
–1	–1	–1	1
1	–1	–1	–1
–1	1	1	1
1	1	1	–1

se separarán los alias de AB de CD y permitirán que los bloques sean ortogonales (esto puede verse desarrollando la matriz **X** como se hizo anteriormente).

En general, suele ser directo el examen de la matriz **X** del modelo reducido que se obtiene de un diseño factorial fraccionado, así como la determinación de cuáles son las corridas que habrán de aumentarse en el diseño original para separar los alias de las interacciones de interés potencial. Además, el impacto de las estrategias específicas para aumentar el diseño puede evaluarse utilizando los resultados generales de los modelos de regresión que se presentan más adelante en este capítulo. Se cuenta también con métodos basados en computadora para construir diseños que pueden ser útiles en el **aumento del diseño** para separar los alias de los efectos. Estos **diseños generados por computadora** se revisarán en el capítulo siguiente.

10-4 PRUEBA DE HIPÓTESIS EN LA REGRESIÓN MÚLTIPLE

En los problemas de regresión lineal múltiple, ciertas pruebas de hipótesis acerca de los parámetros del modelo son una ayuda para medir la utilidad del modelo. En esta sección se describen varios procedimientos de prueba de hipótesis importantes. Estos procedimientos requieren que los errores ε_i del modelo sigan una distribución normal e independiente con media cero y varianza σ^2, lo cual se abrevia $\varepsilon \sim \text{NID}(0, \sigma^2)$. Como resultado de este supuesto, las observaciones y_i tienen una distribución normal e independiente con media $\beta_0 + \Sigma_{j=1}^{k}\beta_j x_{ij}$ y varianza σ^2.

10-4.1 Prueba de significación de la regresión

La prueba de significación de la regresión es un procedimiento para determinar si existe una relación lineal entre la variable de respuesta y y un subconjunto de los regresores $x_1, x_2, ..., x_k$. Las hipótesis apropiadas son

$$\begin{aligned} &H_0: \beta_1 = \beta_2 = \cdots = \beta_k = 0 \\ &H_1: \beta_j \neq 0 \qquad \text{para al menos una } j \end{aligned} \tag{10-20}$$

El rechazo de H_0 de la ecuación 10-20 implica que al menos uno de los regresores $x_1, x_2, ..., x_k$ contribuye de manera significativa al modelo. El procedimiento de prueba incluye un análisis de varianza en el que se

hace la partición de la suma de cuadrados total SS_T en una suma de cuadrados debida al modelo (o a la regresión) y una suma de cuadrados debida a los residuales (o al error), es decir,

$$SS_T = SS_R + SS_E \tag{10-21}$$

Ahora bien, si la hipótesis nula $H_0{:}\beta_1 = \beta_2 = \ldots = \beta_k = 0$ es verdadera, entonces SS_R/σ^2 se distribuye como χ^2_k, donde el número de grados de libertad para χ^2 es igual al número de regresores del modelo. Asimismo, puede demostrarse que SS_E/σ^2 se distribuye como χ^2_{n-k-1} y que SS_E y SS_R son independientes. El procedimiento de prueba para $H_0{:}\beta_1 = \beta_2 = \ldots = \beta_k = 0$ consiste en calcular

$$F_0 = \frac{SS_R / k}{SS_E / (n-k-1)} = \frac{MS_R}{MS_E} \tag{10-22}$$

y en rechazar H_0 si F_0 excede a $F_{\alpha,k,n-k-1}$. De manera alternativa, podría usarse el enfoque del valor P para la prueba de hipótesis y, por lo tanto, rechazar H_0 si el valor P del estadístico F_0 es menor que α. Por lo general la prueba se resume en una tabla del análisis de varianza como la tabla 10-6.

Es sencillo encontrar una fórmula para calcular SS_R. En la ecuación 10-16 se estableció una fórmula para calcular SS_E; es decir,

$$SS_E = \mathbf{y}'\mathbf{y} - \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y}$$

Ahora bien, puesto que $SS_T = \Sigma_{i=1}^n y_i^2 - (\Sigma_{i=1}^n y_i)^2 / n = \mathbf{y}'\mathbf{y} - (\Sigma_{i=1}^n y_i)^2 / n$, la ecuación anterior puede reescribirse como

$$SS_E = \mathbf{y}'\mathbf{y} - \frac{\left(\sum_{i=1}^n y_i\right)^2}{n} - \left[\hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} - \frac{\left(\sum_{i=1}^n y_i\right)^2}{n}\right]$$

o

$$SS_E = SS_T - SS_R$$

Por lo tanto, la suma de cuadrados de regresión es

$$SS_R = \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} - \frac{\left(\sum_{i=1}^n y_i\right)^2}{n} \tag{10-23}$$

mientras la suma de cuadrados del error es

$$SS_E = \mathbf{y}'\mathbf{y} - \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} \tag{10-24}$$

Tabla 10-6 Análisis de varianza de la significación de la regresión en una regresión múltiple

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0
Regresión	SS_R	k	MS_R	MS_R/MS_E
Error o residual	SS_E	$n-k-1$	MS_E	
Total	SS_T	$n-1$		

y la suma de cuadrados total es

$$SS_T = \mathbf{y}'\mathbf{y} - \frac{\left(\sum_{i=1}^{n} y_i\right)^2}{n} \tag{10-25}$$

Estos cálculos casi siempre se realizan con software de regresión. Por ejemplo, en la tabla 10-4 se muestra una parte de la salida de Minitab para el modelo de regresión de la viscosidad del ejemplo 10-1. La sección superior de esta presentación es el análisis de varianza del modelo. La prueba de significación de la regresión en este ejemplo incluye las hipótesis

$$H_0: \beta_1 = \beta_2 = 0$$
$$H_1: \beta_j \neq 0 \quad \text{para al menos una } j$$

El valor P de la tabla 10-4 para el estadístico F (ecuación 10-22) es muy pequeño, por lo que se concluiría que al menos una de las dos variables —la temperatura (x_1) y la velocidad de alimentación (x_2)— tiene un coeficiente de regresión diferente de cero.

En la tabla 10-4 se presenta también el coeficiente de determinación múltiple R^2, donde

$$R^2 = \frac{SS_R}{SS_T} = 1 - \frac{SS_E}{SS_T} \tag{10-26}$$

Como en los experimentos diseñados, R^2 es una medida de la cantidad de reducción en la variabilidad de y que se obtiene al utilizar las variables de regresión $x_1, x_2, ..., x_k$ en el modelo. Sin embargo, como se señaló antes, un valor grande de R^2 no implica necesariamente que el modelo de regresión sea adecuado. Siempre que se agregue una variable al modelo, R^2 se incrementará, independientemente de que la variable adicional sea estadísticamente significativa o no. Por lo tanto, es posible que los modelos que tienen valores grandes de R^2 produzcan predicciones pobres de nuevas observaciones o estimaciones pobres de la respuesta media.

Puesto que R^2 siempre se incrementa cuando se agregan términos al modelo, algunos constructores de modelos de regresión prefieren usar el **estadístico R^2 ajustada** definido como

$$R^2_{\text{ajustada}} = 1 - \frac{SS_E/(n-p)}{SS_T/(n-1)} = 1 - \left(\frac{n-1}{n-p}\right)(1-R^2) \tag{10-27}$$

En general, el estadístico R^2 ajustada no siempre se incrementará cuando se agreguen variables al modelo. De hecho, si se agregan términos innecesarios, el valor de R^2_{ajustada} se decrementará con frecuencia.

Por ejemplo, considere el modelo de regresión de la viscosidad. La R^2 ajustada para el modelo se muestra en la tabla 10-4. Se calcula como

$$\begin{aligned} R^2_{\text{ajustada}} &= 1 - \left(\frac{n-1}{n-p}\right)(1-R^2) \\ &= 1 - \left(\frac{15}{13}\right)(1-0.92697) \\ &= 0.915735 \end{aligned}$$

que está muy cerca de la R^2 ordinaria. Cuando la diferencia entre R^2 y R^2_{ajustada} es considerable, existe un buen riesgo de que se hayan incluido en el modelo términos no significativos.

10-4.2 Pruebas de los coeficientes de regresión individuales y de grupos de coeficientes

Muchas veces el interés se centra en probar hipótesis sobre los coeficientes de regresión individuales. Estas pruebas serían útiles para determinar el valor de cada uno de los regresores del modelo de regresión. Por ejemplo, el modelo podría ser más eficaz con la inclusión de variables adicionales o quizá con la eliminación de una o más de las variables que están ya en el modelo.

Agregar una variable al modelo de regresión ocasiona siempre que la suma de cuadrados de regresión se incremente y que la suma de cuadrados del error se decremente. Es necesario decidir si el incremento de la suma de cuadrados de regresión es suficiente para garantizar el uso de la variable adicional en el modelo. Además, agregar una variable no importante al modelo en realidad puede incrementar el cuadrado medio del error, reduciéndose así la utilidad del modelo.

Las hipótesis para probar la significación de cualquier coeficiente de regresión individual, por ejemplo β_j, son

$$H_0: \beta_j = 0$$
$$H_1: \beta_j \neq 0$$

Si H_0:$\beta_j = 0$ no se rechaza, entonces esto indica que x_j puede eliminarse del modelo. El estadístico de prueba para esta hipótesis es

$$t_0 = \frac{\hat{\beta}_j}{\sqrt{\hat{\sigma}^2 C_{jj}}} \tag{10-28}$$

donde C_{jj} es el elemento de la diagonal de $(\mathbf{X}'\mathbf{X})^{-1}$ correspondiente a $\hat{\beta}_j$. La hipótesis nula H_0:$\beta_j = 0$ se rechaza si $|t_0| > t_{\alpha/2,n-k-1}$. Observe que se trata en realidad de una prueba parcial o marginal, ya que el coeficiente de regresión $\hat{\beta}_j$ depende de todos los demás regresores x_i $(i \neq j)$ que están en el modelo.

Al denominador de la ecuación 10-28, $\sqrt{\hat{\sigma}^2 C_{jj}}$, se le llama con frecuencia **error estándar** (*se*) del coeficiente de regresión $\hat{\beta}_j$. Es decir,

$$se(\hat{\beta}_j) = \sqrt{\hat{\sigma}^2 C_{jj}} \tag{10-29}$$

Por lo tanto, una manera equivalente de escribir el estadístico de prueba de la ecuación 10-28 es

$$t_0 = \frac{\hat{\beta}_j}{se(\hat{\beta}_j)} \tag{10-30}$$

La mayoría de los programas de computadora de regresión proporcionan la prueba t para cada parámetro del modelo. Por ejemplo, considere la tabla 10-4, la cual contiene la salida de Minitab para el ejemplo 10-1. En la sección superior de esta tabla se da la estimación de mínimos cuadrados de cada parámetro, el error estándar, el estadístico t y el valor P correspondiente. Se concluiría que ambas variables, la temperatura y la velocidad de alimentación, contribuyen de manera significativa en el modelo.

También puede examinarse directamente la contribución de una variable particular, por ejemplo x_j, a la suma de cuadrados de regresión, dado que otras x_i variables $(i \neq j)$ están incluidas en el modelo. El procedimiento para hacer esto es la prueba general de la significación de la regresión o, como se denomina con frecuencia, el **método de suma de cuadrados extra.** Este procedimiento también puede usarse para investigar la contribución de un *subconjunto* de los regresores al modelo. Considere el modelo de regresión con k regresores:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

donde $\mathbf{y}$ es $(n \times 1)$, $\mathbf{X}$ es $(n \times p)$, $\boldsymbol{\beta}$ es $(p \times 1)$, $\boldsymbol{\varepsilon}$ es $(n \times 1)$ y $p = k + 1$. Querría determinarse si el subconjunto de regresores $x_1, x_2, ..., x_r$ $(r < k)$ contribuye significativamente al modelo de regresión. Sea que se haga la partición del vector de los coeficientes de regresión de la siguiente manera:

$$\boldsymbol{\beta} = \begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{bmatrix}$$

donde $\boldsymbol{\beta}_1$ es $(r \times 1)$ y $\boldsymbol{\beta}_2$ es $[(p - r) \times 1]$. Quieren probarse las hipótesis

$$H_0: \boldsymbol{\beta}_1 = \mathbf{0}$$
$$H_1: \boldsymbol{\beta}_1 \neq \mathbf{0} \qquad (10\text{-}31)$$

El modelo puede escribirse como

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon} \qquad (10\text{-}32)$$

donde $\mathbf{X}_1$ representa las columnas de $\mathbf{X}$ asociadas con $\boldsymbol{\beta}_1$ y $\mathbf{X}_2$ representa las columnas de $\mathbf{X}$ asociadas con $\boldsymbol{\beta}_2$.

Para el **modelo completo** (incluyendo tanto a $\boldsymbol{\beta}_1$ como a $\boldsymbol{\beta}_2$) se sabe que $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$. Además, la suma de cuadrados de regresión para todas las variables incluyendo la ordenada al origen es

$$SS_R(\boldsymbol{\beta}) = \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} \quad (p \text{ grados de libertad})$$

y

$$MS_E = \frac{\mathbf{y}'\mathbf{y} - \hat{\boldsymbol{\beta}}\mathbf{X}'\mathbf{y}}{n - p}$$

A $SS_R(\boldsymbol{\beta})$ se le llama la suma de cuadrados de regresión debida a $\boldsymbol{\beta}$. Para encontrar la contribución de los términos en $\boldsymbol{\beta}_1$ a la regresión, se ajusta el modelo suponiendo que la hipótesis nula H_0:$\boldsymbol{\beta}_1 = \mathbf{0}$ es verdadera. El **modelo reducido** se encuentra a partir de la ecuación 10-32 con $\boldsymbol{\beta}_1 = \mathbf{0}$:

$$\mathbf{y} = \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon} \qquad (10\text{-}33)$$

El estimador de mínimos cuadrados de $\boldsymbol{\beta}_2$ es $\hat{\boldsymbol{\beta}}_2 = (\mathbf{X}'_2\mathbf{X}_2)^{-1}\mathbf{X}'_2\mathbf{y}$, y

$$SS_R(\boldsymbol{\beta}_2) = \hat{\boldsymbol{\beta}}_2'\mathbf{X}_2'\mathbf{y} \quad (p - r \text{ grados de libertad}) \qquad (10\text{-}34)$$

La suma de cuadrados de regresión debida a $\boldsymbol{\beta}_1$ dado que $\boldsymbol{\beta}_2$ está ya en el modelo es

$$SS_R(\boldsymbol{\beta}_1|\boldsymbol{\beta}_2) = SS_R(\boldsymbol{\beta}) - SS_R(\boldsymbol{\beta}_2) \qquad (10\text{-}35)$$

Esta suma de cuadrados tiene r grados de libertad. Es la "suma de cuadrados extra" debida a $\boldsymbol{\beta}_1$. Observe que $SS_R(\boldsymbol{\beta}_1|\boldsymbol{\beta}_2)$ es el incremento en la suma de cuadrados de regresión debido a la inclusión de las variables $x_1, x_2, ..., x_r$ en el modelo.

Ahora bien, $SS_R(\boldsymbol{\beta}_1|\boldsymbol{\beta}_2)$ es independiente de MS_E, y la hipótesis nula $\boldsymbol{\beta}_1 = \mathbf{0}$ puede probarse con el estadístico

$$F_0 = \frac{SS_R(\boldsymbol{\beta}_1|\boldsymbol{\beta}_2)/r}{MS_E} \qquad (10\text{-}36)$$

Si $F_0 > F_{\alpha,r,n-p}$, se rechaza H_0, y se concluye que al menos uno de los parámetros en $\boldsymbol{\beta}_1$ es diferente de cero y, por consiguiente, al menos una de las variables $x_1, x_2, ..., x_r$ en $\mathbf{X}_1$ contribuye significativamente al modelo de regresión. Algunos autores llaman a la prueba de la ecuación 10-36 la **prueba F parcial.**

La prueba F parcial es muy útil. Puede usarse para medir la contribución de x_j como si fuera la última variable que se agregó al modelo, calculando

$$SS_R(\beta_j | \beta_0, \beta_1, \ldots, \beta_{j-1}, \beta_{j+1}, \ldots, \beta_k)$$

Éste es el incremento en la suma de cuadrados de regresión debido a que se agrega x_j a un modelo que ya contiene a $x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_k$. Observe que la prueba F parcial de una sola variable x_j es equivalente a la prueba t de la ecuación 10-28. Sin embargo, la prueba F parcial es un procedimiento más general por cuanto puede medir el efecto de conjuntos de variables.

EJEMPLO 10-6 ··

Considere los datos de la viscosidad del ejemplo 10-1. Suponga que se quiere investigar la contribución de la variable x_2 (velocidad de alimentación) al modelo. Es decir, las hipótesis que quieren probarse son

$$H_0: \beta_2 = 0$$
$$H_1: \beta_2 \neq 0$$

Esto requerirá la suma de cuadrados extra debida a β_2, o

$$\begin{aligned} SS_R(\beta_2 | \beta_1, \beta_0) &= SS_R(\beta_0, \beta_1, \beta_2) - SS_R(\beta_0, \beta_1) \\ &= SS_R(\beta_1, \beta_2 | \beta_0) - SS_R(\beta_2 | \beta_0) \end{aligned}$$

Entonces, por la tabla 10-4, donde se probó la significación de la regresión, se tiene

$$SS_R(\beta_1, \beta_2 | \beta_0) = 44{,}157.1$$

a la que se llamó en la tabla la suma de cuadrados del modelo. Esta suma de cuadrados tiene dos grados de libertad.

El modelo reducido es

$$y = \beta_0 + \beta_1 x_1 + \varepsilon$$

El ajuste de mínimos cuadrados de este modelo es

$$\hat{y} = 1652.3955 + 7.6397 x_1$$

y la suma de cuadrados de regresión para este modelo (con un grado de libertad) es

$$SS_R(\beta_1 | \beta_0) = 40{,}840.8$$

Observe que $SS_R(\beta_1 | \beta_0)$ se muestra en la parte inferior de la salida de Minitab de la tabla 10-4 bajo el encabezado "Seq SS". Por lo tanto,

$$\begin{aligned} SS_R(\beta_2 | \beta_0, \beta_1) &= 44{,}157.1 - 40{,}840.8 \\ &= 3316.3 \end{aligned}$$

con $2 - 1 = 1$ grado de libertad. Éste es el incremento en la suma de cuadrados de regresión que resulta de agregar x_2 a un modelo que contenía ya a x_1, y se muestra en la parte inferior de la salida de Minitab en la tabla 10-4. Para probar $H_0: \beta_2 = 0$, por el estadístico de prueba se obtiene

$$F_0 = \frac{SS_R(\beta_2 | \beta_0, \beta_1)/1}{MS_E} = \frac{3316.3/1}{267.604} = 12.3926$$

Observe que en el denominador de F_0 se usa MS_E del modelo completo (tabla 10-4). Entonces, puesto que $F_{0.05,1,13} = 1.67$, se rechazaría $H_0: \beta_2 = 0$ y se concluiría que x_2 (velocidad de alimentación) contribuye significativamente al modelo.

Debido a que esta prueba F parcial incluye un solo regresor, es equivalente a la prueba t porque el cuadrado de una variable aleatoria t con ν grados de libertad es una variable aleatoria F con 1 y ν grados de libertad. Para ver esto, observe, por la tabla 10-4, que el estadístico t para H_0:$\beta_2 = 0$ dio como resultado $t_0 = 3.5203$ y que $t_0^2 = (3.5203)^2 = 12.3925 \simeq F_0$.

10-5 INTERVALOS DE CONFIANZA EN REGRESIONES MÚLTIPLES

Con frecuencia es necesario construir estimaciones de intervalos de confianza para los coeficientes de regresión $\{\beta_j\}$ y para otras cantidades de interés del modelo de regresión. El desarrollo de un procedimiento para obtener estos intervalos de confianza requiere suponer que los errores $\{\varepsilon_i\}$ tienen una distribución normal e independiente con media cero y varianza σ^2, el mismo supuesto que se estableció en la sección sobre la prueba de hipótesis de la sección 10-4.

10-5.1 Intervalos de confianza para los coeficientes de regresión individuales

Puesto que el estimador de mínimos cuadrados $\hat{\boldsymbol{\beta}}$ es una combinación lineal de las observaciones, se sigue que $\hat{\boldsymbol{\beta}}$ tiene una distribución normal con vector medio $\boldsymbol{\beta}$ y matriz de covarianza $\sigma^2(\mathbf{X}'\mathbf{X})^{-1}$. Entonces cada uno de los estadísticos

$$\frac{\hat{\beta}_j - \beta_j}{\sqrt{\hat{\sigma}^2 C_{jj}}} \qquad j = 0, 1, \ldots, k \tag{10-37}$$

se distribuye como t con $n-p$ grados de libertad, donde C_{jj} es el elemento (jj)-ésimo de la matriz $(\mathbf{X}'\mathbf{X})^{-1}$, y $\hat{\sigma}^2$ es la estimación de la varianza del error, obtenida con la ecuación 10-17. Por lo tanto, un intervalo de confianza de $100(1-\alpha)$ por ciento para el coeficiente de regresión β_j, $j = 0, 1, \ldots, k$, es

$$\hat{\beta}_j - t_{\alpha/2,n-p}\sqrt{\hat{\sigma}^2 C_{jj}} \le \beta_j \le \hat{\beta}_j + t_{\alpha/2,n-p}\sqrt{\hat{\sigma}^2 C_{jj}} \tag{10-38}$$

Observe que este intervalo de confianza también podría escribirse como

$$\hat{\beta}_j - t_{\alpha/2,n-p}se(\hat{\beta}_j) \le \beta_j \le \hat{\beta}_j + t_{\alpha/2,n-p}se(\hat{\beta}_j)$$

ya que $se(\hat{\beta}_j) = \sqrt{\hat{\sigma}^2 C_{jj}}$.

EJEMPLO 10-7

Se construirá un intervalo de confianza de 95% para el parámetro β_1 del ejemplo 10-1. Ahora bien, $\hat{\beta}_1 = 7.62129$, y puesto que $\hat{\sigma}^2 = 267.604$ y $C_{11} = 1.429184 \times 10^{-3}$, se encuentra que

$$\hat{\beta}_1 - t_{0.025,13}\sqrt{\hat{\sigma}^2 C_{11}} \le \beta_1 \le \hat{\beta}_1 + t_{0.025,13}\sqrt{\hat{\sigma}^2 C_{11}}$$

$$7.62129 - 2.16\sqrt{(267.604)(1.429184\times 10^{-3})} \le \beta_1$$

$$\le 7.62129 + 2.16\sqrt{(267.604)(1.429184\times 10^{-3})}$$

$$7.62129 - 2.16(0.6184) \le \beta_1 \le 7.62129 + 2.16(0.6184)$$

y el intervalo de confianza de 95% para β_1 es

$$6.2855 \le \beta_1 \le 8.9570$$

10-5.2 Intervalo de confianza para la respuesta media

También puede obtenerse un intervalo de confianza para la respuesta media en un punto particular, por ejemplo, $x_{01}, x_{02}, ..., x_{0k}$. Se define primero el vector

$$\mathbf{x}_0 = \begin{bmatrix} 1 \\ x_{01} \\ x_{02} \\ \vdots \\ x_{0k} \end{bmatrix}$$

La respuesta media en este punto es

$$\mu_{y|\mathbf{x}_0} = \beta_0 + \beta_1 x_{01} + \beta_2 x_{02} + \cdots + \beta_k x_{0k} = \mathbf{x}_0' \boldsymbol{\beta}$$

La respuesta media estimada en este punto es

$$\hat{y}(\mathbf{x}_0) = \mathbf{x}_0' \hat{\boldsymbol{\beta}} \tag{10-39}$$

Este estimador es insesgado, ya que $E[\hat{y}(\mathbf{x}_0)] = E(\mathbf{x}_0'\hat{\boldsymbol{\beta}}) = \mathbf{x}_0'\boldsymbol{\beta} = \mu_{y|\mathbf{x}_0}$, y la varianza de $\hat{y}(\mathbf{x}_0)$ es

$$V[\hat{y}(\mathbf{x}_0)] = \sigma^2 \mathbf{x}_0'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_0 \tag{10-40}$$

Por lo tanto, un intervalo de confianza de $100(1-\alpha)$ por ciento para la respuesta media en el punto $x_{01}, x_{02}, ..., x_{0k}$ es

$$\hat{y}(\mathbf{x}_0) - t_{\alpha/2,n-p}\sqrt{\hat{\sigma}^2 \mathbf{x}_0'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_0} \le \mu_{y|\mathbf{x}_0} \le \hat{y}(\mathbf{x}_0) + t_{\alpha/2,n-p}\sqrt{\hat{\sigma}^2 \mathbf{x}_0'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_0} \tag{10-41}$$

10-6 PREDICCIÓN DE NUEVAS OBSERVACIONES DE LA RESPUESTA

Es posible usar un modelo de regresión para predecir observaciones futuras de la respuesta y que corresponden a valores particulares de los regresores, por ejemplo $x_{01}, x_{02}, ..., x_{0k}$. Si $\mathbf{x}'_0 = [1, x_{01}, x_{02}, ..., x_{0k}]$, entonces una estimación puntual de la observación futura y_0 en el punto $x_{01}, x_{02}, ..., x_{0k}$ se calcula con la ecuación 10-39:

$$\hat{y}(\mathbf{x}_0) = \mathbf{x}_0' \hat{\boldsymbol{\beta}}$$

Un **intervalo de predicción** de $100(1-\alpha)$ por ciento para esta observación futura es

$$\hat{y}(\mathbf{x}_0) - t_{\alpha/2,n-p}\sqrt{\hat{\sigma}^2 (1 + \mathbf{x}_0'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_0)} \le y_0 \le \hat{y}(\mathbf{x}_0) + t_{\alpha/2,n-p}\sqrt{\hat{\sigma}^2 (1 + \mathbf{x}_0'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_0)} \tag{10-42}$$

Cuando se predicen nuevas observaciones y se estima la respuesta media en un punto dado $x_{01}, x_{02}, ..., x_{0k}$, es necesario tener cuidado para no hacer una extrapolación fuera de la región que contiene las observaciones originales. Es muy posible que un modelo que se ajuste bien en la región de los datos originales deje de hacerlo fuera de esa región.

10-7 DIAGNÓSTICOS DEL MODELO DE REGRESIÓN

Como se destacó en los experimentos diseñados, la **verificación de la adecuación del modelo** es una parte importante en el procedimiento del análisis de datos. Es de igual importancia en la construcción de modelos de regresión y, como se ilustró en el ejemplo 10-1, en un modelo de regresión deberán examinarse

siempre las **gráficas de los residuales** que se usaron en los experimentos diseñados. En general, siempre es necesario: 1) examinar el modelo ajustado para asegurarse de que proporciona una aproximación adecuada del verdadero sistema y 2) verificar que no se infringe ninguno de los supuestos de la regresión de mínimos cuadrados. El modelo de regresión probablemente producirá resultados pobres o equivocados a menos que sea un ajuste adecuado.

Además de las gráficas de los residuales, existen otros diagnósticos del modelo que con frecuencia son útiles en la regresión. En esta sección se presenta un breve resumen de estos procedimientos. Para análisis más completos, ver Montgomery y Peck [82] y Myers [84].

10-7.1 Residuales escalados y PRESS

Residuales estandarizados y studentizados

Muchos constructores de modelos prefieren trabajar con **residuales escalados** en lugar de los residuales de mínimos cuadrados ordinarios. Estos residuales escalados transmiten con frecuencia más información que los residuales ordinarios.

Un tipo de residual escalado es el **residual estandarizado:**

$$d_i = \frac{e_i}{\hat{\sigma}} \qquad i = 1, 2, \ldots, n \tag{10-43}$$

donde por lo general se usa $\hat{\sigma} = \sqrt{MS_E}$ en los cálculos. Estos residuales estandarizados tienen media cero y varianza aproximadamente unitaria; por consiguiente, son muy útiles para buscar **puntos atípicos.** La mayoría de los residuales estandarizados deberán localizarse en el intervalo $-3 \leq d_i \leq 3$, y cualquier observación con un residual estandarizado que esté fuera de este intervalo es potencialmente inusual con respecto a su respuesta observada. Estos puntos atípicos deberán examinarse con atención, ya que pueden representar algo tan simple como un error al registrar los datos o algo que sea motivo de mayor preocupación, como una región del espacio del regresor, donde el modelo ajustado es una aproximación pobre de la verdadera superficie de respuesta.

El proceso de estandarización de la ecuación 10-43 escala los residuales al dividirlos por su desviación estándar promedio aproximada. En algunos conjuntos de datos, los residuales pueden tener desviaciones estándar que difieren considerablemente. A continuación se presenta una escalación que toma en consideración esta situación.

El vector de los valores ajustados $\hat{y}_i$ que corresponden a los valores observados y_i es

$$\begin{aligned} \hat{\mathbf{y}} &= \mathbf{X}\hat{\boldsymbol{\beta}} \\ &= \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \\ &= \mathbf{H}\mathbf{y} \end{aligned} \tag{10-44}$$

A la matriz $n \times n$, $\mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ se le llama generalmente la matriz "gorro" porque mapea el vector de los valores observados en un vector de los valores ajustados. La matriz gorro y sus propiedades desempeñan un papel central en el análisis de regresión.

Los residuales del modelo ajustado pueden escribirse convenientemente en la notación matricial como

$$\mathbf{e} = \mathbf{y} - \hat{\mathbf{y}}$$

y resulta que la matriz de covarianza de los residuales es

$$\text{Cov}(\mathbf{e}) = \sigma^2(\mathbf{I} - \mathbf{H}) \tag{10-45}$$

La matriz $\mathbf{I} - \mathbf{H}$ no es por lo general diagonal, por lo que los residuales tienen varianzas diferentes y están correlacionados.

Por lo tanto, la varianza del residual i-ésimo es

$$V(e_i) = \sigma^2(1-h_{ii}) \tag{10-46}$$

donde h_{ii} es el elemento i-ésimo de la diagonal de **H**. Puesto que $0 \leq h_{ii} \leq 1$, al utilizar el cuadrado medio residual MS_E para estimar la varianza de los residuales en realidad se está sobreestimando $V(e_i)$. Además, puesto que h_{ii} es una medida de localización del punto i-ésimo en el espacio x, la varianza de e_i depende de dónde esté el punto x_i. En general, los residuales situados cerca del centro del espacio x tienen varianzas más grandes que los residuales situados en lugares más apartados. Las violaciones de los supuestos del modelo son más probables en los puntos remotos, y estas violaciones pueden ser difíciles de detectar por la inspección de e_i (o d_i) porque sus residuales serán por lo general más pequeños.

Se recomienda tomar en consideración esta desigualdad de la varianza cuando se escalen los residuales. Se sugiere graficar los **residuales studentizados:**

$$r_i = \frac{e_i}{\sqrt{\hat{\sigma}^2(1-h_{ii})}} \qquad i = 1, 2, \ldots, n \tag{10-47}$$

con $\hat{\sigma}^2 = MS_E$ en lugar de e_i (o d_i). Los residuales studentizados tienen varianza constante $V(r_i) = 1$ independientemente de la localización de $\mathbf{x}_i$ cuando la forma del modelo es correcta. En muchas situaciones la varianza de los residuales se estabiliza, en particular para conjuntos de datos grandes. En estos casos puede haber poca diferencia entre los residuales estandarizados y los studentizados. Por lo tanto, los residuales estandarizados y studentizados transmiten con frecuencia información equivalente. Sin embargo, ya que cualquier punto con un residual grande y una h_{ii} grande tiene una influencia potencialmente considerable sobre el ajuste de mínimos cuadrados, suele recomendarse el examen de los residuales studentizados. En la tabla 10-3 se presentan las diagonales gorro h_{ii} y los residuales studentizados para el modelo de regresión de la viscosidad del ejemplo 10-1.

Residuales PRESS

La suma de cuadrados del error de predicción (PRESS, del inglés Prediction Error Sum of Squares) proporciona una útil escalación de los residuales. Para calcular la PRESS se selecciona una observación, por ejemplo la i. Se ajusta el modelo de regresión a las $n-1$ observaciones restantes y se usa esta ecuación para predecir la observación que se apartó y_i. Al denotar este valor predicho $\hat{y}_{(i)}$, puede encontrarse el error de predicción del punto i como $e_{(i)} = y_i = \hat{y}_{(i)}$. Al error de predicción suele llamársele el residual PRESS i-ésimo. Este procedimiento se repite para cada observación $i = 1, 2, \ldots, n$, produciéndose un conjunto de n residuales PRESS $e_{(1)}, e_{(2)}, \ldots, e_{(n)}$. Entonces el estadístico PRESS se define como la suma de cuadrados de los n residuales PRESS como en

$$\text{PRESS} = \sum_{i=1}^{n} e_{(i)}^2 = \sum_{i=1}^{n} [y_i - \hat{y}_{(i)}]^2 \tag{10-48}$$

Por lo tanto, la PRESS utiliza cada subconjunto posible de $n-1$ observaciones como un conjunto de datos de estimación, y se utiliza una observación a la vez para formar un conjunto de datos de predicción.

Inicialmente, parecería que para calcular la PRESS es necesario ajustar n regresiones diferentes. Sin embargo, la PRESS puede calcularse a partir de los resultados de un solo ajuste de mínimos cuadrados a las n observaciones totales. Resulta que el residual PRESS i-ésimo es

$$e_{(i)} = \frac{e_i}{1-h_{ii}} \tag{10-49}$$

Por lo tanto, ya que la PRESS es tan sólo la suma de cuadrados de los residuales PRESS, una fórmula de cálculo simple es

$$\text{PRESS} = \sum_{i=1}^{n} \left(\frac{e_i}{1-h_{ii}}\right)^2 \tag{10-50}$$

Por la ecuación 10-49 es sencillo ver que el residual PRESS es sólo el residual ordinario ponderado de acuerdo con los elementos de la diagonal de la matriz gorro h_{ii}. Los puntos de los datos para los que h_{ii} es grande tendrán residuales PRESS grandes. Estas observaciones serán por lo general puntos de **alta influencia.** En general, una diferencia grande entre el residual ordinario y los residuales PRESS indicará un punto donde el modelo se ajusta bien a los datos, pero un modelo construido sin dicho punto producirá predicciones pobres. En la siguiente sección se estudiarán otras medidas de influencia.

Por último, cabe señalar que la PRESS puede usarse para calcular una R^2 aproximada de predicción, por ejemplo

$$R^2_{\text{Predicción}} = 1 - \frac{\text{PRESS}}{S_{yy}} \qquad (10\text{-}51)$$

Este estadístico ofrece cierto indicio de la capacidad predictiva del modelo de regresión. Para el modelo de regresión de la viscosidad del ejemplo 10-1, los residuales PRESS pueden calcularse utilizando los residuales ordinarios y el valor de h_{ii} encontrado en la tabla 10-3. El valor correspondiente del estadístico PRESS es PRESS = 5207.7. Entonces

$$\begin{aligned} R^2_{\text{Predicción}} &= 1 - \frac{\text{PRESS}}{S_{yy}} \\ &= 1 - \frac{5207.7}{47{,}635.9} \\ &= 0.8907 \end{aligned}$$

Por lo tanto, podría esperarse que este modelo "explique" cerca de 89% de la variabilidad al predecir nuevas observaciones, en comparación con el aproximadamente 93% de la variabilidad en los datos originales que explica el ajuste de mínimos cuadrados. La capacidad predictiva global del modelo basado en este criterio parece ser muy satisfactoria.

R-student

Es común considerar al residual studentizado r_i comentado antes como el diagnóstico de un punto atípico. Se acostumbra usar MS_E como una estimación de σ^2 en el cálculo de r_i. Se hace referencia a este enfoque como la escalación interna del residual, ya que MS_E es una estimación de σ^2 generada internamente que se obtiene del ajuste del modelo a las n observaciones. Otro enfoque sería usar una estimación de σ^2 basada en un conjunto de datos en el que se elimina la observación i-ésima. La estimación de σ^2 así obtenida se denota por $S^2_{(i)}$. Puede demostrarse que

$$S^2_{(i)} = \frac{(n-p)MS_E - e_i^2/(1-h_{ii})}{n-p-1} \qquad (10\text{-}52)$$

La estimación de σ^2 de la ecuación 10-52 se usa en lugar de MS_E para producir un residual studentizado externamente, al que es común llamar R-student, dado por

$$t_i = \frac{e_i}{\sqrt{S^2_{(i)}(1-h_{ii})}} \qquad i = 1, 2, \ldots, n \qquad (10\text{-}53)$$

En muchas situaciones habrá una ligera diferencia entre t_i y el residual studentizado r_i. Sin embargo, si la observación i-ésima es influyente, entonces $S^2_{(i)}$ puede diferir significativamente de MS_E, y por lo tanto la R-student será más sensible a este punto. Además, bajo los supuestos usuales, t_i tiene una distribución t_{n-p-1}. Por lo tanto, la R-student ofrece un procedimiento más formal para detectar puntos atípicos a través de la prueba de hipótesis. En la tabla 10-3 se muestran los valores de la R-student para el modelo de regresión de la viscosidad del ejemplo 10-1. Ninguno de esos valores es inusualmente grande.

10-7.2 Diagnósticos de influencia

En ocasiones se encuentra que un subconjunto pequeño de los datos ejerce una influencia desproporcionada sobre el modelo de regresión ajustado. Es decir, las estimaciones o predicciones de los parámetros pueden depender más del subconjunto influyente que de la mayoría de los datos. Sería conveniente localizar estos puntos influyentes y valorar su impacto en el modelo. Si estos puntos influyentes son valores "malos", deberán eliminarse. Por otra parte, quizá no haya nada malo con estos puntos. Pero si controlan propiedades clave del modelo, sería deseable saberlo, ya que podría afectar el uso del modelo. En esta sección se describen e ilustran algunas medidas útiles de influencia.

Puntos de acción de palanca

La localización de los puntos en el espacio x es importante para determinar las propiedades del modelo. En particular, las observaciones apartadas tienen potencialmente acciones o brazos de palanca desproporcionados sobre las estimaciones de los parámetros, los valores predichos y los estadísticos de resumen usuales.

La matriz gorro $\mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ es muy útil para identificar las observaciones influyentes. Como ya se señaló, $\mathbf{H}$ determina las varianzas y covarianzas de $\hat{\mathbf{y}}$ y $\mathbf{e}$, ya que $V(\hat{\mathbf{y}}) = \sigma^2\mathbf{H}$ y $V(\mathbf{e}) = \sigma^2(\mathbf{I} - \mathbf{H})$. Los elementos h_{ij} de $\mathbf{H}$ pueden interpretarse como la cantidad de acción de palanca ejercida por y_j sobre $\hat{y}_i$. Por lo tanto, la inspección de los elementos de $\mathbf{H}$ puede revelar puntos que son potencialmente influyentes en virtud de su localización en el espacio x. La atención suele centrarse en los elementos de la diagonal h_{ii}. Puesto que $\Sigma_{i=1}^{n} h_{ii} = \text{rango}(\mathbf{H}) = \text{rango}(\mathbf{X}) = p$, el tamaño promedio de los elementos de la diagonal de la matriz $\mathbf{H}$ es p/n. Como guía aproximada, entonces, si un elemento h_{ii} de la diagonal es mayor que $2p/n$, la observación i es un punto con acción de palanca alta. Para aplicar lo anterior al modelo de la viscosidad del ejemplo 10-1, observe que $2p/n = 2(3)/16 = 0.375$. En la tabla 10-3 se dan las diagonales gorro h_{ii} para el modelo de primer orden; puesto que ninguna de las h_{ii} excede 0.375, se concluiría que no hay puntos de acción de palanca en estos datos.

Influencia sobre los coeficientes de regresión

Las diagonales gorro identificarán los puntos potencialmente influyentes debido a su localización en el espacio x. Es deseable considerar la localización del punto y la variable de respuesta cuando se mide la influencia. Cook [32a, b] ha sugerido el uso de una medida $\hat{\boldsymbol{\beta}}$ del cuadrado de la distancia entre la estimación de mínimos cuadrados basada en todos los n puntos y la estimación obtenida al eliminar el punto i, por ejemplo $\hat{\boldsymbol{\beta}}_{(i)}$. Esta medida de la distancia puede expresarse como

$$D_i = \frac{(\hat{\boldsymbol{\beta}}_{(i)} - \hat{\boldsymbol{\beta}})'\mathbf{X}'\mathbf{X}(\hat{\boldsymbol{\beta}}_{(i)} - \hat{\boldsymbol{\beta}})}{pMS_E} \qquad i = 1, 2, \ldots, n \tag{10-54}$$

Un valor de referencia razonable para D_i es la unidad. Es decir, en general las observaciones para las que $D_i > 1$ se consideran influyentes.

El estadístico D_i se calcula en realidad a partir de

$$D_i = \frac{r_i^2}{p}\frac{V[\hat{y}(x_i)]}{V(e_i)} = \frac{r_i^2}{p}\frac{h_{ii}}{(1-h_{ii})} \qquad i = 1, 2, \ldots, n \tag{10-55}$$

Observe que, aparte de la constante p, D_i es el producto del cuadrado del residual studentizado i-ésimo y $h_{ii}/(1-h_{ii})$. Puede demostrarse que este cociente es la distancia del vector $\mathbf{x}_i$ al centroide de los datos restantes. Por lo tanto, D_i está compuesto por un componente que refleja la medida en que el modelo ajusta

la observación i-ésima y_i y un componente que mide qué tan alejado está ese punto del resto de los datos. Cualquiera de los componentes (o ambos) puede contribuir a un valor grande de D_i.

En la tabla 10-3 se muestran los valores de D_i para el ajuste del modelo de regresión a los datos de la viscosidad del ejemplo 10-1. Ninguno de estos valores de D_i excede 1, por lo que no hay evidencia sólida de observaciones influyentes en estos datos.

10-8 PRUEBA DE FALTA DE AJUSTE

En la sección 6-6 se indicó cómo agregar puntos centrales a un diseño factorial 2^k le permite al experimentador obtener una estimación del error experimental puro. Esto permite hacer la partición de la suma de cuadrados de los residuales SS_E en dos componentes; es decir,

$$SS_E = SS_{PE} + SS_{LOF}$$

donde SS_{PE} es la suma de cuadrados debida al error puro y SS_{LOF} es la suma de cuadrados debida a la falta de ajuste.

Puede presentarse un desarrollo general de esta partición en el contexto de un modelo de regresión. Suponga que se tienen n_i observaciones de la respuesta en el nivel i-ésimo de los regresores $\mathbf{x}_i$, $i = 1, 2, ..., m$. Sea que y_{ij} denota la observación j-ésima de la respuesta en $\mathbf{x}_i$, i = 1, 2, ..., m y $j = 1, 2, ..., n_i$. Hay $n = \Sigma_{i=1}^m n_i$ observaciones en total. El residual (ij)-ésimo puede escribirse como

$$y_{ij} - \hat{y}_i = (y_{ij} - \bar{y}_i) + (\bar{y}_i - \hat{y}_i) \tag{10-56}$$

donde $\bar{y}_i$ es el promedio de las n_i observaciones en $\mathbf{x}_i$. Al elevar al cuadrado ambos miembros de la ecuación 10-56 y hacer la operación suma sobre i y j se obtiene

$$\sum_{i=1}^{m}\sum_{j=1}^{n_i} (y_{ij} - \hat{y}_i)^2 = \sum_{i=1}^{m}\sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2 + \sum_{i=1}^{m} n_i(\bar{y}_i - \hat{y}_i)^2 \tag{10-57}$$

El primer miembro de la ecuación 10-57 es la suma de cuadrados de los residuales ordinaria. Los dos componentes del segundo miembro miden el error puro y la falta de ajuste. Se observa que la suma de cuadrados del error puro

$$SS_{PE} = \sum_{i=1}^{m}\sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2 \tag{10-58}$$

se obtiene calculando la suma de cuadrados corregida de las observaciones repetidas en cada nivel de $\mathbf{x}$ y haciendo después la agrupación en los m niveles de $\mathbf{x}$. Si se satisface el supuesto de la varianza constante, ésta es una medida **independiente del modelo** del error puro, ya que para calcular SS_{PE} sólo se usa la variabilidad de las y en cada nivel $\mathbf{x}_i$. Puesto que hay $n_i - 1$ grados de libertad del error puro en cada nivel $\mathbf{x}_i$, el número total de grados de libertad asociados con la suma de cuadrados del error puro es

$$\sum_{i=1}^{m} (n_i - 1) = n - m \tag{10-59}$$

La suma de cuadrados de la falta de ajuste

$$SS_{LOF} = \sum_{i=1}^{m} n_i(\bar{y}_i - \hat{y}_i)^2 \tag{10-60}$$

es una suma ponderada de los cuadrados de las desviaciones entre la respuesta media $\bar{y}_i$ en cada nivel $\mathbf{x}_i$ y el valor ajustado correspondiente. Si los valores ajustados $\hat{y}_i$ están cerca de las respuestas promedio $\bar{y}_i$ correspondientes, entonces hay un fuerte indicio de que la función de regresión es lineal. Si las $\hat{y}_i$ se desvían mucho de las $\bar{y}_i$, entonces es probable que la función de regresión no sea lineal. Hay $m-p$ grados de libertad asociados con SS_{LOF} porque hay m niveles de **x**, y se pierden p grados de libertad porque deben estimarse p parámetros para el modelo. En lo que a los cálculos se refiere, por lo general SS_{LOF} se obtiene restando SS_{PE} de SS_E.

El estadístico de prueba para la falta de ajuste es

$$F_0 = \frac{SS_{LOF}/(m-p)}{SS_{PE}/(n-m)} = \frac{MS_{LOF}}{MS_{PE}} \tag{10-61}$$

El valor esperado de MS_{PE} es σ^2, y el valor esperado de MS_{LOF} es

$$E(MS_{LOF}) = \sigma^2 + \frac{\sum_{i=1}^{m} n_i \left[E(y_i) - \beta_0 - \sum_{j=1}^{k} \beta_j x_{ij} \right]^2}{m-2} \tag{10-62}$$

Si la verdadera función de regresión es lineal, entonces $E(y_i) = \beta_0 + \Sigma_{j=1}^{k} \beta_j x_{ij}$, y el segundo término de la ecuación 10-62 es cero, dando como resultado $E(MS_{LOF}) = \sigma^2$. Sin embargo, si la verdadera función de regresión no es lineal, entonces $E(y_i) \neq \beta_0 + \Sigma_{j=1}^{k} \beta_j x_{ij}$ y $E(MS_{LOF}) > \sigma^2$. Además, si la verdadera función de regresión es lineal, entonces el estadístico F_0 sigue la distribución $F_{m-p,n-m}$. Por lo tanto, para probar la falta de ajuste, se calcularía el estadístico de prueba F_0 y se concluiría que la función de regresión no es lineal si $F_0 > F_{\alpha,\, m-p,n-m}$.

Es sencillo incorporar este procedimiento de prueba en el análisis de varianza. Si se concluye que la función de regresión no es lineal, entonces el modelo tentativo habrá de abandonarse y deberán hacerse intentos para encontrar una ecuación más apropiada. De manera alternativa, si F_0 no excede $F_{\alpha,m-p,n-m}$, no existe evidencia sólida de falta de ajuste y MS_{PE} y MS_{LOF} se combinan con frecuencia para estimar σ^2. El ejemplo 6-6 es una ilustración muy completa de este procedimiento, donde las réplicas de las corridas son puntos centrales de un diseño factorial 2^2.

10-9 PROBLEMAS

10-1. La resistencia a la tensión de un producto de papel se relaciona con la cantidad de madera dura en la pulpa. Se producen 10 muestras en la planta piloto y los datos obtenidos se presentan en la siguiente tabla.

Resistencia	Porcentaje de madera dura	Resistencia	Porcentaje de madera dura
160	10	181	20
171	15	188	25
175	15	193	25
182	20	195	28
184	20	200	30

a) Ajustar un modelo de regresión lineal que relacione la resistencia con el porcentaje de madera dura.

b) Probar el modelo del inciso *a* para la significación de la regresión.

c) Encontrar un intervalo de confianza de 95% para el parámetro β_1.

10-2. En una planta se destila aire líquido para producir oxígeno, nitrógeno y argón. Se piensa que el porcentaje de impurezas en el oxígeno se relaciona linealmente con la cantidad de impurezas en el aire, medida por el

"conteo de contaminación" en partes por millón (ppm). Una muestra de los datos de operación de la planta se presenta a continuación:

Pureza (%)	93.3	92.0	92.4	91.7	94.0	94.6	93.6	93.1	93.2	92.9	92.2	91.3	90.1	91.6	91.9
Conteo de contaminación (ppm)	1.10	1.45	1.36	1.59	1.08	0.75	1.20	0.99	0.83	1.22	1.47	1.81	2.03	1.75	1.68

a) Ajustar un modelo de regresión lineal a los datos.
b) Probar la significación de la regresión.
c) Encontrar un intervalo de confianza de 95% para β_1.

10-3. Graficar los residuales del problema 10-1 y comentar la adecuación del modelo.

10-4. Graficar los residuales del problema 10-2 y comentar la adecuación del modelo.

10-5. Utilizando los resultados del problema 10-1, probar el modelo de regresión para la falta de ajuste.

10-6. Se realizó un estudio sobre el desgaste y de un cojinete y su relación con x_1 = viscosidad del aceite y x_2 = carga. Se obtuvieron los siguientes datos:

y	x_1	x_2
193	1.6	851
230	15.5	816
172	22.0	1058
91	43.0	1201
113	33.0	1357
125	40.0	1115

a) Ajustar un modelo de regresión lineal múltiple a los datos.
b) Probar la significación de la regresión.
c) Calcular el estadístico t para cada parámetro del modelo. ¿Qué conclusiones pueden sacarse?

10-7. Se piensa que la potencia al freno desarrollada por el motor de un automóvil en un dinamómetro es una función de la rapidez del motor en revoluciones por minuto (rpm), el octanaje del combustible y la compresión del motor. Se llevó a cabo un experimento en el laboratorio y los datos colectados fueron:

Potencia al freno	rpm	Octanaje	Compresión
225	2000	90	100
212	1800	94	95
229	2400	88	110
222	1900	91	96
219	1600	86	100
278	2500	96	110
246	3000	94	98
237	3200	90	100
233	2800	88	105
224	3400	86	97
223	1800	90	100
230	2500	89	104

a) Ajustar un modelo de regresión múltiple a estos datos.
b) Probar la significación de la regresión. ¿Qué conclusiones pueden sacarse?
c) Con base en las pruebas t, ¿son necesarios los tres regresores en el modelo?

10-8. Analizar los residuales del modelo de regresión del problema 10-7. Comentar la adecuación del modelo.

10-9. El rendimiento de un proceso químico se relaciona con la concentración del reactivo y la temperatura de operación. Se realiza un experimento con los siguientes resultados:

Rendimiento	Concentración	Temperatura
81	1.00	150
89	1.00	180
83	2.00	150
91	2.00	180
79	1.00	150
87	1.00	180
84	2.00	150
90	2.00	180

a) Suponga que quiere ajustarse un modelo de los efectos principales a estos datos. Establecer la matriz **X'X** utilizando los datos exactamente como aparecen en la tabla.
b) ¿La matriz que se obtuvo en el inciso *a* es diagonal? Comentar la respuesta.
c) Suponga que el modelo se escribe en términos de las variables codificadas "usuales"

$$x_1 = \frac{\text{Concentración} - 1.5}{0.5} \quad x_2 = \frac{\text{Temperatura} - 165}{15}$$

Establecer la matriz **X'X** para el modelo en términos de estas variables codificadas. ¿Esta matriz es diagonal? Comentar la respuesta.

d) Definir un nuevo conjunto de variables codificadas

$$x_1 = \frac{\text{Concentración} - 1.0}{1.0} \quad x_2 = \frac{\text{Temperatura} - 150}{30}$$

Establecer la matriz **X'X** para el modelo en términos de este conjunto de variables codificadas. ¿Esta matriz es diagonal? Comentar la respuesta.

e) Resumir lo que se haya aprendido acerca de la codificación de variables con este problema.

10-10. Considere el experimento factorial 2^4 del ejemplo 6-2. Suponga que falta la última observación. Volver a analizar los datos y sacar conclusiones. ¿Cómo se comparan estas conclusiones con las del ejemplo original?

10-11. Considere el experimento factorial 2^4 del ejemplo 6-2. Suponga que faltan las dos últimas observaciones. Volver a analizar los datos y sacar conclusiones. ¿Cuál es el resultado de la comparación de estas conclusiones con las del ejemplo original?

10-12. Dados los datos siguientes, ajustar el modelo de regresión polinomial de segundo orden

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{12} x_1 x_2 + \varepsilon$$

y	x_1	x_2
26	1.0	1.0
24	1.0	1.0
175	1.5	4.0
160	1.5	4.0
163	1.5	4.0
55	0.5	2.0
62	1.5	2.0
100	0.5	3.0
26	1.0	1.5
30	0.5	1.5
70	1.0	2.5
71	1.5	2.5

Después de que se haya ajustado el modelo, probar la significación de la regresión.

10-13. *a*) Considere el modelo de regresión cuadrático del problema 10-12. Calcular los estadísticos t de cada uno de los parámetros del modelo y comentar las conclusiones a que se llega a partir de estas cantidades.

b) Usar el método de la suma de cuadrados extra para evaluar el valor de los términos cuadráticos x_1^2, x_2^2 y x_1x_2 del modelo.

10-14. ***Relación entre el análisis de varianza y el análisis de regresión.*** Cualquier modelo del análisis de varianza puede expresarse en términos del modelo lineal general $\mathbf{y} = \mathbf{x}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, donde la matriz $\mathbf{X}$ se compone de ceros y unos. Demostrar que el modelo con un solo factor $y_{ij} = \mu + \tau_i + \varepsilon_{ij}$, $i = 1, 2, 3, j = 1, 2, 3, 4$ puede escribirse en la forma del modelo lineal general. Después

a) Escribir las ecuaciones normales $(\mathbf{X'X})\hat{\boldsymbol{\beta}} = \mathbf{X'y}$ y compararlas con las ecuaciones normales que se encontraron en el capítulo 3 para este modelo.

b) Encontrar el rango de $\mathbf{X'X}$. ¿Es posible obtener $(\mathbf{X'X})^{-1}$?

c) Suponga que se elimina la primera ecuación normal y se agrega la restricción $\Sigma_{i=1}^{3} n\hat{\tau}_i = 0$. ¿Tiene solución el sistema de ecuaciones resultante? De ser así, encontrarla. Hallar la suma de cuadrados de regresión $\hat{\boldsymbol{\beta}}'\mathbf{X'y}$ y compararla con la suma de cuadrados de los tratamientos del modelo con un solo factor.

10-15. Suponga que se está haciendo el ajuste de una línea recta y se desea hacer la varianza de $\hat{\beta}_1$ tan pequeña como sea posible. Al trabajar con la restricción de un número par de puntos experimentales, ¿dónde deberán colocarse estos puntos para minimizar $V(\hat{\beta}_1)$? (*Nota*: usar el diseño que se pide en este ejercicio con *sumo* cuidado, ya que, aun cuando minimiza $V(\hat{\beta}_1)$, tiene propiedades indeseables; ver, por ejemplo, Myers y Montgomery [85a]. Únicamente si se tiene una *gran seguridad* de que la verdadera relación funcional es lineal deberá considerarse el uso de este diseño.)

10-16. ***Mínimos cuadrados ponderados.*** Suponga que se está ajustando la línea recta $y = \beta_0 + \beta_1 x + \varepsilon$, pero la varianza de las y depende ahora del nivel de x; es decir,

$$V(y \mid x_i) = \sigma_i^2 = \frac{\sigma^2}{w_i} \qquad i = 1, 2, \ldots, n$$

donde las w_i son constantes desconocidas, llamadas con frecuencia ponderaciones. Demostrar que si se eligen las estimaciones de los coeficientes de regresión para minimizar la suma de cuadrados de los errores ponderados dada por $\sum_{i=1}^{n} w_i(y_i - \beta_0 - \beta_1 x_i)^2$, las ecuaciones normales de mínimos cuadrados resultantes son

$$\hat{\beta}_0 \sum_{i=1}^{n} w_i + \hat{\beta}_1 \sum_{i=1}^{n} w_i x_i = \sum_{i=1}^{n} w_i y_i$$

$$\hat{\beta}_0 \sum_{i=1}^{n} w_i x_i + \hat{\beta}_1 \sum_{i=1}^{n} w_i x_i^2 = \sum_{i=1}^{n} w_i x_i y_i$$

10-17. Considere el diseño 2_{IV}^{4-1} analizado en el ejemplo 10-5.

a) Suponga que se opta por aumentar el diseño con la corrida única seleccionada en ese ejemplo. Encontrar las varianzas y las covarianzas de los coeficientes de regresión del modelo (ignorando los bloques):

$$y=\beta_0+\beta_1x_1+\beta_2x_2+\beta_3x_3+\beta_4x_4$$
$$+\beta_{12}x_1x_2+\beta_{34}x_3x_4+\varepsilon$$

b) ¿Hay otras corridas de la fracción alterna que separarían los alias *AB* de *CD*?

c) Suponga que el diseño se aumenta con las cuatro corridas sugeridas en el ejemplo 10-5. Encontrar las varianzas y las covarianzas de los coeficientes de regresión (ignorando los bloques) para el modelo del inciso *a*.

d) Considerando los incisos *a* y *c*, ¿qué estrategia de aumento se preferiría y por qué?

10-18. Considere un diseño 2_{III}^{7-4}. Suponga que después de correr el experimento, los efectos observados más grandes son $A + BD$, $B + AD$ y $D + AB$. Quiere aumentarse el diseño original con un grupo de cuatro corridas para separar los alias de estos efectos.

a) ¿Cuáles son las cuatro corridas que se harían?

b) Encontrar las varianzas y las covarianzas de los coeficientes de regresión del modelo

$$y=\beta_0+\beta_1x_1+\beta_2x_2+\beta_4x_4+\beta_{12}x_1x_2+\beta_{14}x_1x_4$$
$$+\beta_{24}x_2x_4+\varepsilon$$

c) ¿Es posible separar los alias de estos efectos con menos de cuatro corridas adicionales?

11 Métodos de superficies de respuesta y otros enfoques para la optimización de procesos

11-1 INTRODUCCIÓN A LA METODOLOGÍA DE SUPERFICIES DE RESPUESTA

La metodología de superficies de respuesta, o **MSR**, es una colección de técnicas matemáticas y estadísticas útiles en el modelado y el análisis de problemas en los que una respuesta de interés recibe la influencia de diversas variables y donde el objetivo es optimizar esta respuesta. Por ejemplo, suponga que un ingeniero químico quiere encontrar los niveles de temperatura (x_1) y presión (x_2) que maximicen el rendimiento (y) de un proceso. El rendimiento del proceso es una función de los niveles de la temperatura y la presión, por ejemplo,

$$y = f(x_1, x_2) + \varepsilon$$

donde ε representa el ruido o error observado en la respuesta y. Si la respuesta esperada se denota por $E(y) = f(x_1, x_2) = \eta$, entonces a la superficie representada por

$$\eta = f(x_1, x_2)$$

se le llama **superficie de respuesta.**

Por lo general la superficie de respuesta se representa gráficamente como en la figura 11-1, donde η se grafica contra los niveles de x_1 y x_2. Se han visto ya gráficas de superficie de respuesta como ésta, particularmente en los capítulos sobre diseños factoriales. Para ayudar a visualizar la forma de una superficie de respuesta, con frecuencia se grafican los contornos de la superficie de respuesta, como se muestra en la figura 11-2. En la gráfica de contorno se trazan las líneas de respuesta constante en el plano x_1, x_2. Cada contorno corresponde a una altura particular de la superficie de respuesta. También se ha visto antes la utilidad de las gráficas de contorno.

En la mayoría de los problemas MSR, la forma de la relación entre la respuesta y las variables independientes es desconocida. Por lo tanto, el primer paso de la MSR es encontrar una aproximación adecuada de la verdadera relación funcional entre y y el conjunto de variables independientes. Por lo general se emplea un polinomio de orden inferior en alguna región de las variables independientes. Si la respuesta está bien modelada por una función lineal de las variables independientes, entonces la función de aproximación es el **modelo de primer orden**

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \varepsilon \qquad (11\text{-}1)$$

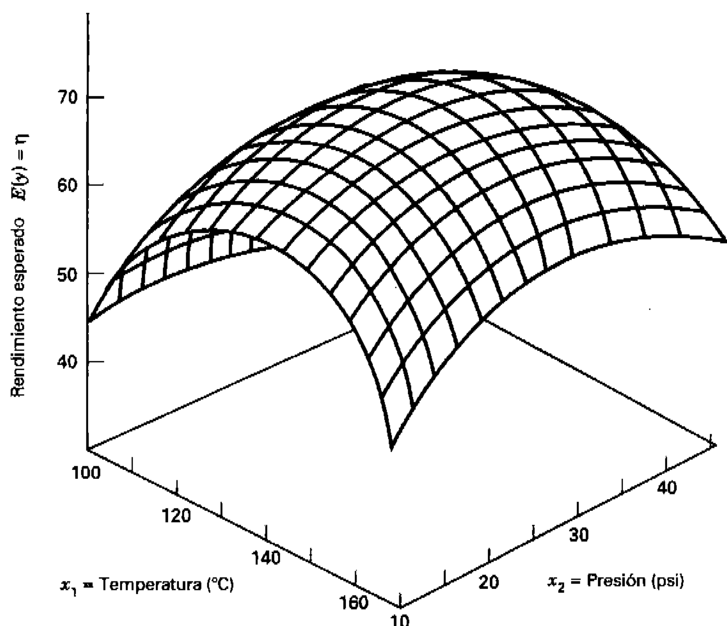


Figura 11-1 Superficie de respuesta tridimensional donde se indica el rendimiento esperado (η) como una función de la temperatura (x_1) y la presión (x_2).

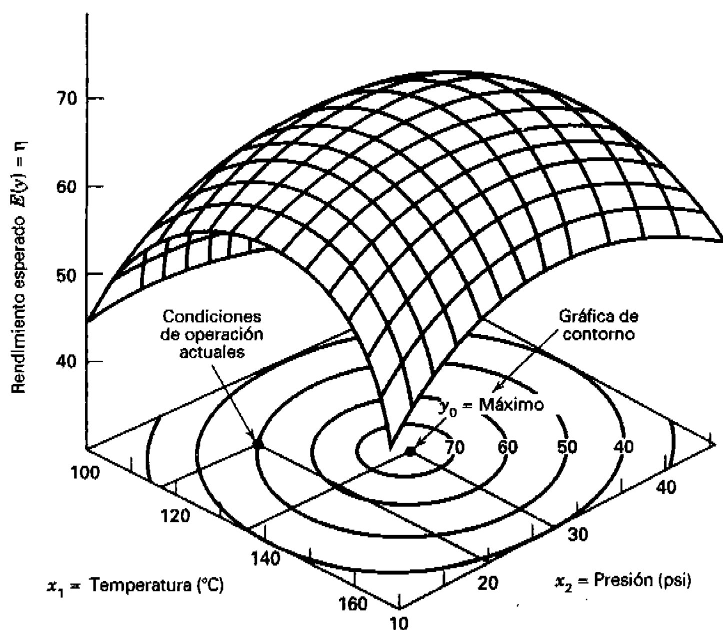


Figura 11-2 Gráfica de contorno de una superficie de respuesta.

Si hay curvatura en el sistema, entonces debe usarse un polinomio de orden superior, tal como el **modelo de segundo orden**

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum_{i<j}\sum \beta_{ij} x_i x_j + \varepsilon \tag{11-2}$$

En casi todos los problemas MSR se usa uno de estos modelos, o ambos. Desde luego, es probable que un modelo polinomial sea una aproximación razonable de la verdadera relación funcional en el espacio completo de las variables independientes, pero para una región relativamente pequeña suelen funcionar bastante bien.

El método de mínimos cuadrados, estudiado en el capítulo 10, se usa para estimar los parámetros de los polinomios de aproximación. Después se realiza el análisis de la superficie de respuesta utilizando la superficie ajustada. Si la superficie ajustada es una aproximación adecuada de la verdadera función de la respuesta, entonces el análisis de la superficie ajustada será un equivalente aproximado del análisis del sistema real. Los parámetros del modelo pueden estimarse de manera más eficiente cuando se emplean los diseños experimentales apropiados para recolectar los datos. Los diseños para ajustar superficies de respuesta se denominan **diseños de superficie de respuesta**. Estos diseños se revisan en la sección 11-4.

La MSR es un **procedimiento secuencial**. Muchas veces, cuando se está en un punto de la superficie de respuesta que está apartado del óptimo, como en el caso de las condiciones de operación actuales de la figura 11-3, el sistema presenta una curvatura moderada y el modelo de primer orden será apropiado. El objetivo en este caso es llevar al experimentador de manera rápida y eficiente por la trayectoria del mejoramiento hasta la vecindad general del **óptimo**. Una vez que se ha encontrado la región del óptimo, puede emplearse un modelo más elaborado, como el de segundo orden, y llevarse a cabo un análisis para localizar el óptimo. En la figura 11-3 se puede ver que el análisis de una superficie de respuesta puede considerarse como "el ascenso a una colina", donde la cima de ésta representa el punto de la respuesta máxima. Si el verdadero óptimo es un punto de respuesta mínima, entonces la situación puede considerarse como "el descenso a un valle".

El objetivo último de la MSR es determinar las condiciones de operación óptimas del sistema o determinar una región del espacio de los factores en la que se satisfagan los requerimientos de operación.

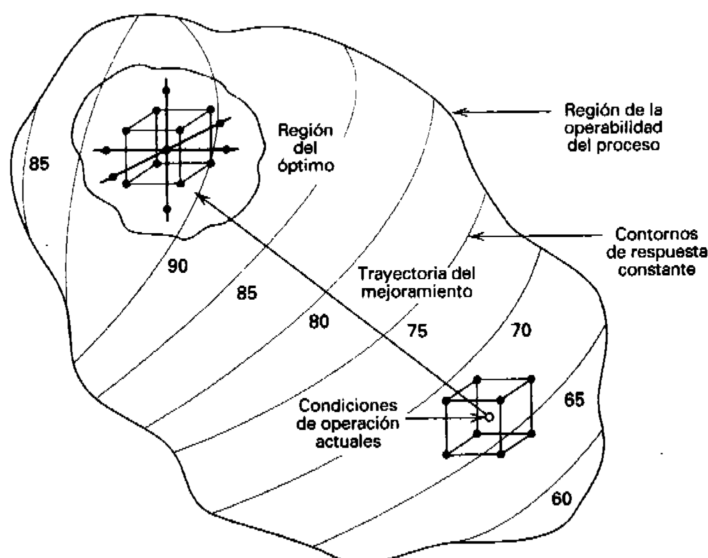


Figura 11-3 El carácter secuencial de la MSR.

Análisis más detallados de la MSR se encuentran en Myers y Montgomery [85a], Khuri y Cornell [67] y Box y Draper [16b].

11-2 MÉTODO DEL ASCENSO MÁS PRONUNCIADO

Frecuentemente la estimación inicial de las condiciones de operación óptimas del sistema estarán lejos del óptimo real. En tales circunstancias, el objetivo del experimentador es pasar con rapidez a la vecindad general del óptimo. Para ello desea usarse un procedimiento experimental económico y eficiente. Cuando se está muy lejos del óptimo, por lo general se supone que un modelo de primer orden es una aproximación adecuada de la verdadera superficie en una región pequeña de las x.

El **método del ascenso más pronunciado** es un procedimiento para moverse secuencialmente sobre la trayectoria del ascenso más pronunciado, es decir, en la dirección del incremento máximo de la respuesta. Desde luego, si lo que se pretende es una minimización, entonces esta técnica se llama **método del descenso más pronunciado.** El modelo ajustado de primer orden es

$$\hat{y} = \hat{\beta}_0 + \sum_{i=1}^{k} \hat{\beta}_i x_i \tag{11-3}$$

y la superficie de respuesta de primer orden, es decir, los contornos de $\hat{y}$, es una serie de líneas paralelas como las que se muestran en la figura 11-4. La dirección del ascenso más pronunciado es aquella en la que $\hat{y}$ se incrementa con mayor rapidez. Esta dirección es paralela a la normal de la superficie de respuesta ajustada. Por lo general se toma como la **trayectoria del ascenso más pronunciado** a la recta que pasa por el centro de la región de interés y que es normal a la superficie ajustada. Por lo tanto, los pasos sobre la

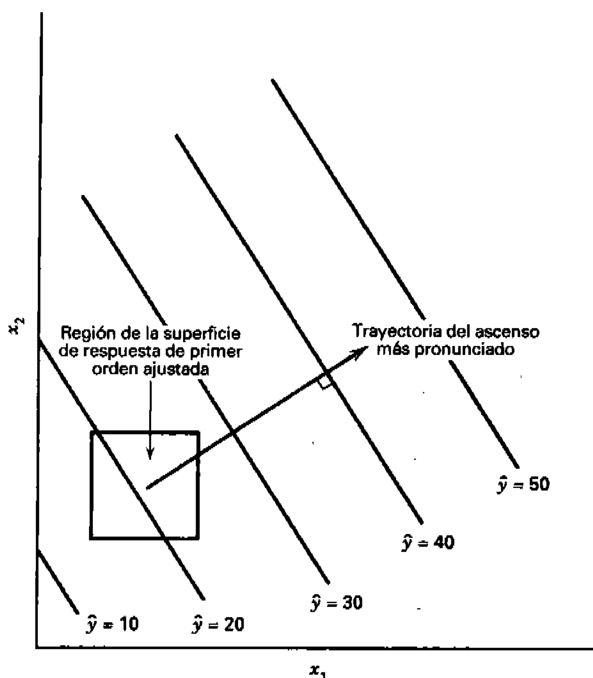


Figura 11-4 Superficie de respuesta de primer orden y trayectoria del ascenso más pronunciado.

trayectoria son proporcionales a los coeficientes de regresión $\{\hat{\beta}_i\}$. El tamaño real del paso lo determina el experimentador con base en el conocimiento del proceso o de otras consideraciones prácticas.

Se conducen experimentos sobre la trayectoria del ascenso más pronunciado hasta que deja de observarse un incremento adicional en la respuesta. Entonces puede ajustarse un nuevo modelo de primer orden, determinarse una nueva trayectoria del ascenso más pronunciado y el procedimiento continúa. En última instancia, el experimentador llegará a la vecindad del óptimo. En general, la falta de ajuste del modelo de primer orden indica que se ha llegado a ella. En este momento se realizan experimentos adicionales para obtener una estimación más precisa del óptimo.

EJEMPLO 11-1

Un ingeniero químico está interesado en determinar las condiciones de operación que maximizan el rendimiento de un proceso. Dos variables controlables influyen en el rendimiento del proceso: el tiempo de reacción y la temperatura de reacción. El ingeniero opera actualmente el proceso con un tiempo de reacción de 35 minutos y una temperatura de 155°F, que dan como resultado rendimientos de cerca de 40%. Puesto que es improbable que esta región contenga el óptimo, el ingeniero ajusta un modelo de primer orden y aplica el método del ascenso más pronunciado.

El ingeniero decide que la región de exploración para ajustar el modelo de primer orden deberá ser (30, 40) minutos de tiempo de reacción y (150, 160)°F. Para simplificar los cálculos, las variables independientes se codificarán en el intervalo usual (–1, 1). Por lo tanto, si ξ_1 denota la **variable natural** tiempo y ξ_2 la **variable natural** temperatura, entonces las **variables codificadas** son

$$x_1 = \frac{\xi_1 - 35}{5} \quad \text{y} \quad x_2 = \frac{\xi_2 - 155}{5}$$

El diseño experimental se muestra en la tabla 11-1. Observe que el diseño usado para recabar estos datos es un factorial 2^2 aumentado con cinco puntos centrales. Las réplicas del centro se usan para estimar el error experimental y permitir la verificación de la adecuación del modelo de primer orden. Además, el diseño está centrado alrededor de las condiciones de operación actuales del proceso.

Es posible ajustar un modelo de primer orden a estos datos por el procedimiento de mínimos cuadrados. Aplicando los métodos para diseños de dos niveles se obtiene el siguiente modelo en las variables codificadas:

$$\hat{y} = 40.44 + 0.775x_1 + 0.325x_2$$

Tabla 11-1 Datos del proceso para ajustar el modelo de primer orden

Variables naturales		Variables codificadas		Respuesta
ξ_1	ξ_2	x_1	x_2	y
30	150	–1	–1	39.3
30	160	–1	1	40.0
40	150	1	–1	40.9
40	160	1	1	41.5
35	155	0	0	40.3
35	155	0	0	40.5
35	155	0	0	40.7
35	155	0	0	40.2
35	155	0	0	40.6

Antes de explorar a lo largo de la trayectoria del ascenso más pronunciado, deberá investigarse la adecuación del modelo de primer orden. El diseño 2^2 con puntos centrales permite al experimentador

1. Obtener una estimación del error.
2. Verificar las interacciones (o términos de productos cruzados) del modelo.
3. Verificar los efectos cuadráticos (curvatura).

Las réplicas del centro pueden usarse para calcular una estimación del error de la siguiente manera:

$$\begin{aligned}\hat{\sigma}^2 &= \frac{(40.3)^2 + (40.5)^2 + (40.7)^2 + (40.2)^2 + (40.6)^2 - (202.3)^2/5}{4} \\ &= 0.0430\end{aligned}$$

En el modelo de primer orden se supone que las variables x_1 y x_2 tienen un **efecto aditivo** sobre la respuesta. La interacción entre las variables se representaría por el coeficiente β_{12} del término de un producto cruzado x_1x_2 sumado al modelo. La estimación de mínimos cuadrados de este coeficiente es simplemente la mitad del efecto de la interacción que se calcula como en un diseño factorial 2^2 ordinario, o

$$\begin{aligned}\hat{\beta}_{12} &= \tfrac{1}{4}[(1\times 39.3)+(1\times 41.5)+(-1\times 40.0)+(-1\times 40.9)] \\ &= \tfrac{1}{4}(-0.1) \\ &= -0.025\end{aligned}$$

La suma de cuadrados de la interacción con un solo grado de libertad es

$$\begin{aligned}SS_{\text{Interacción}} &= \frac{(-0.1)^2}{4} \\ &= 0.0025\end{aligned}$$

Al comparar $SS_{\text{Interacción}}$ con $\hat{\sigma}^2$ se obtiene el estadístico para la falta de ajuste

$$\begin{aligned}F &= \frac{SS_{\text{Interacción}}}{\hat{\sigma}^2} \\ &= \frac{0.0025}{0.0430} \\ &= 0.058\end{aligned}$$

que es pequeño, lo cual indica que la interacción es insignificante.

Otra verificación de la adecuación del modelo de línea recta se obtiene aplicando la verificación del efecto de curvatura cuadrática pura de la sección 6-6. Recuerde que ésta consiste en comparar la respuesta promedio en los cuatro puntos de la porción factorial del diseño, por ejemplo $\bar{y}_F = 40.425$, con la respuesta promedio en el centro del diseño, por ejemplo $\bar{y}_C = 40.46$. Si existe curvatura cuadrática en la verdadera función de la respuesta, entonces $\bar{y}_F - \bar{y}_C$ es una medida de esta curvatura. Si β_{11} y β_{22} son los coeficientes de los términos "cuadráticos puros" x_1^2 y x_2^2, entonces $\bar{y}_F - \bar{y}_C$ es una estimación de $\beta_{11} + \beta_{22}$. En el ejemplo tratado aquí, una estimación del término cuadrático puro es

$$\begin{aligned}\hat{\beta}_{11} + \hat{\beta}_{22} &= \bar{y}_F - \bar{y}_C \\ &= 40.425 - 40.46 \\ &= -0.035\end{aligned}$$

Tabla 11-2 Análisis de varianza del modelo de primer orden

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0	Valor P
Modelo (β_1, β_2)	2.8250	2	1.4125	47.83	0.0002
Residual	0.1772	6			
(Interacción)	(0.0025)	1	0.0025	0.058	0.8215
(Cuadrático puro)	(0.0027)	1	0.0027	0.063	0.8142
(Error puro)	(0.1720)	4	0.0430		
Total	3.0022	8			

La suma de cuadrados con un solo grado de libertad asociada con la hipótesis nula, $H_0{:}\beta_{11} + \beta_{22} = 0$, es

$$SS_{\text{Cuadrática pura}} = \frac{n_F n_C (\bar{y}_F - \bar{y}_C)^2}{n_F + n_C}$$
$$= \frac{(4)(5)(-0.035)^2}{4+5}$$
$$= 0.0027$$

donde n_F y n_C son el número de puntos de la porción factorial y el número de puntos centrales, respectivamente. Puesto que

$$F = \frac{SS_{\text{Cuadrática pura}}}{\hat{\sigma}^2}$$
$$= \frac{0.0027}{0.0430}$$
$$= 0.063$$

es pequeño, no hay indicios de un efecto cuadrático puro.

En la tabla 11-2 se resume el análisis de varianza de este modelo. Las verificaciones de la interacción y la curvatura no son significativas, mientras que la prueba F de la regresión global es significativa. Además, el error estándar de $\hat{\beta}_1$ y $\hat{\beta}_2$ es

$$se(\hat{\beta}_i) = \sqrt{\frac{MS_E}{4}} = \sqrt{\frac{\hat{\sigma}^2}{4}} = \sqrt{\frac{0.0430}{4}} = 0.10 \qquad i = 1,\ 2$$

Ambos coeficientes de regresión $\hat{\beta}_1$ y $\hat{\beta}_2$ son grandes en comparación con sus errores estándar. En este punto no hay razón para cuestionar la adecuación del modelo de primer orden.

Para apartarse del centro del diseño —el punto ($x_1 = 0, x_2 = 0$)— sobre la trayectoria del ascenso más pronunciado, se haría un movimiento de 0.775 unidades en la dirección x_1 por cada 0.325 unidades en la dirección x_2. Por lo tanto, la trayectoria del ascenso más pronunciado pasa por el punto ($x_1 = 0, x_2 = 0$) y tiene pendiente 0.325/0.775. El ingeniero decide usar 5 minutos de tiempo de reacción como tamaño básico del paso. Al utilizar la relación entre ξ_1 y x_1, se observa que 5 minutos de tiempo de reacción es equivalente a un paso en la variable *codificada* x_1 de $\Delta x_1 = 1$. Por lo tanto, los pasos sobre la trayectoria del ascenso más pronunciado son $\Delta x_1 = 1.0000$ y $\Delta x_2 = (0.325/0.775)\ \Delta x_1 = 0.42$.

El ingeniero calcula puntos sobre esta trayectoria y observa los rendimientos en los mismos hasta que se nota un decremento en la respuesta. En la tabla 11-3 se muestran los resultados tanto en variables codificadas como naturales. Aun cuando la manipulación matemática de las variables codificadas es más sencilla, deben usarse las variables naturales cuando se corre el proceso. En la figura 11-5 se grafica el

Tabla 11-3 Experimento del ascenso más pronunciado para el ejemplo 11-1

Pasos	Variables codificadas x_1	x_2	Variables naturales ξ_1	ξ_2	Respuesta y
Origen	0	0	35	155	
Δ	1.00	0.42	5	2	
Origen + Δ	1.00	0.42	40	157	41.0
Origen + 2Δ	2.00	0.84	45	159	42.9
Origen + 3Δ	3.00	1.26	50	161	47.1
Origen + 4Δ	4.00	1.68	55	163	49.7
Origen + 5Δ	5.00	2.10	60	165	53.8
Origen + 6Δ	6.00	2.52	65	167	59.9
Origen + 7Δ	7.00	2.94	70	169	65.0
Origen + 8Δ	8.00	3.36	75	171	70.4
Origen + 9Δ	9.00	3.78	80	173	77.6
Origen + 10Δ	10.00	4.20	85	175	80.3
Origen + 11Δ	11.00	4.62	90	179	76.2
Origen + 12Δ	12.00	5.04	95	181	75.1

rendimiento en cada paso de la trayectoria del ascenso más pronunciado. Se observan incrementos de la respuesta hasta el décimo paso; sin embargo, todos los pasos después de este punto resultan en un decremento del rendimiento. Por lo tanto, deberá ajustarse otro modelo de primer orden en la vecindad general del punto ($\xi_1 = 85$, $\xi_2 = 175$).

Se ajusta un nuevo modelo de primer orden alrededor del punto ($\xi_1 = 85$, $\xi_2 = 175$). La región de exploración para ξ_1 es [80, 90] y para ξ_2 es [170, 180]. Por lo tanto, las variables codificadas son

$$x_1 = \frac{\xi_1 - 85}{5} \quad \text{y} \quad x_2 = \frac{\xi_2 - 175}{5}$$

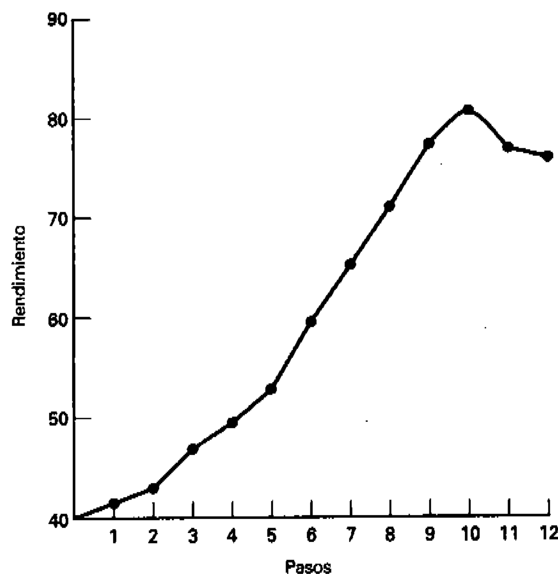


Figura 11-5 Rendimiento contra pasos sobre la trayectoria del ascenso más pronunciado para el ejemplo 11-1.

Tabla 11-4 Datos para el segundo modelo de primer orden

Variables naturales		Variables codificadas		Respuesta
ξ_1	ξ_2	x_1	x_2	y
80	170	–1	–1	76.5
80	180	–1	1	77.0
90	170	1	–1	78.0
90	180	1	1	79.5
85	175	0	0	79.9
85	175	0	0	80.3
85	175	0	0	80.0
85	175	0	0	79.7
85	175	0	0	79.8

De nueva cuenta se usa un diseño 2^2 con cinco puntos centrales. El diseño experimental se muestra en la tabla 11-4.

El ajuste del modelo de primer orden a las variables codificadas de la tabla 11-4 es

$$\hat{y} = 78.97 + 1.00x_1 + 0.50x_2$$

En la tabla 11-5 se presenta el análisis de varianza de este modelo, incluyendo las verificaciones de la interacción y del término cuadrático puro. Las verificaciones de la interacción y del término cuadrático puro implican que el modelo de primer orden no es una aproximación adecuada. Esta curvatura en la verdadera superficie puede indicar que el experimentador se encuentra cerca del óptimo. En este punto es necesario hacer análisis adicionales para localizar el óptimo con mayor precisión.

. . . .

Por el ejemplo 11-1 se observa que la *trayectoria del ascenso más pronunciado es proporcional a los signos y magnitudes de los coeficientes de regresión* del modelo ajustado de primer orden

$$\hat{y} = \hat{\beta}_0 + \sum_{i=1}^{k} \hat{\beta}_i x_i$$

Es sencillo dar un algoritmo general para determinar las coordenadas de un punto sobre la trayectoria del ascenso más pronunciado. Suponga que el punto $x_1 = x_2 = \ldots = x_k = 0$ es la base o punto origen. Entonces

1. Se elige el tamaño del paso en una de las variables del proceso, por ejemplo Δx_j. En general, se seleccionaría la variable de la que se tenga mayor información, o se seleccionaría la variable que tiene el coeficiente de regresión absoluto $|\hat{\beta}_j|$ más grande.

Tabla 11-5 Análisis de varianza del segundo modelo de primer orden

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0	Valor P
Regresión	5.00	2			
Residual	11.1200	6			
(Interacción)	(0.2500)	1	0.2500	4.72	0.0955
(Cuadrático puro)	(10.6580)	1	10.6580	201.09	0.0001
(Error puro)	(0.2120)	4	0.0530		
Total	16.1200	8			

2. El tamaño del paso de las otras variables es

$$\Delta x_i = \frac{\hat{\beta}_i}{\hat{\beta}_j / \Delta x_j} \qquad i = 1, 2, \ldots, k; \qquad i \neq j$$

3. Se convierten las Δx_i de variables codificadas a variables naturales.

Para ilustrar, considere la trayectoria del ascenso más pronunciado calculada en el ejemplo 11-1. Puesto que x_1 tiene el coeficiente de regresión más grande, se selecciona el tiempo de reacción como la variable del paso 1 del procedimiento anterior. Cinco minutos de tiempo de reacción es el tamaño del paso (con base en el conocimiento del proceso). En términos de las variables codificadas, éste es $\Delta x_1 = 1.0$. Por lo tanto, por el lineamiento 2, el tamaño del paso de la temperatura es

$$\Delta x_2 = \frac{\hat{\beta}_2}{\hat{\beta}_1 / \Delta x_1} = \frac{0.325}{(0.775/1.0)} = 0.42$$

Para convertir los tamaños de los pasos codificados ($\Delta x_1 = 1.0$ y $\Delta x_2 = 0.42$) a las unidades naturales de tiempo y temperatura, se usan las relaciones

$$\Delta x_1 = \frac{\Delta \xi_1}{5} \qquad \text{y} \qquad \Delta x_2 = \frac{\Delta \xi_2}{5}$$

que dan como resultado

$$\Delta \xi_1 = \Delta x_1(5) = 1.0(5) = 5 \text{ min}$$

y

$$\Delta \xi_2 = \Delta x_2(5) = 0.42(5) = 2°\text{F}$$

11-3 ANÁLISIS DE UNA SUPERFICIE DE RESPUESTA DE SEGUNDO ORDEN

Cuando el experimentador se encuentra relativamente cerca del óptimo, por lo general se requiere un modelo que incorpore la curvatura para aproximar la respuesta. En la mayoría de los casos, el modelo de segundo orden

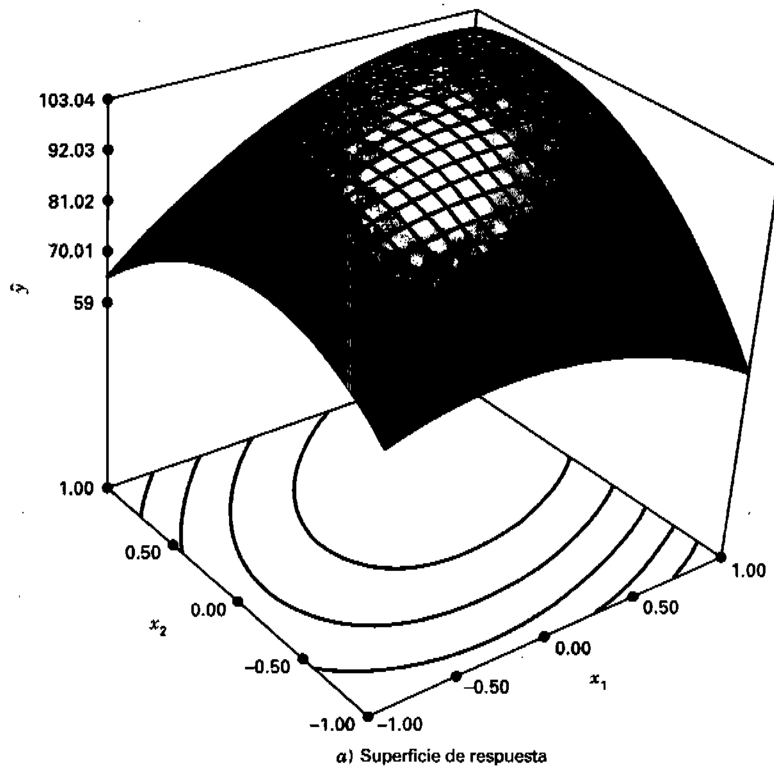
$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum_{i<j}\sum \beta_{ij} x_i x_j + \varepsilon \qquad (11\text{-}4)$$

es adecuado. En esta sección se indicará cómo usar este modelo ajustado para encontrar el conjunto óptimo de condiciones de operación para las x, así como para caracterizar la naturaleza de la superficie de respuesta.

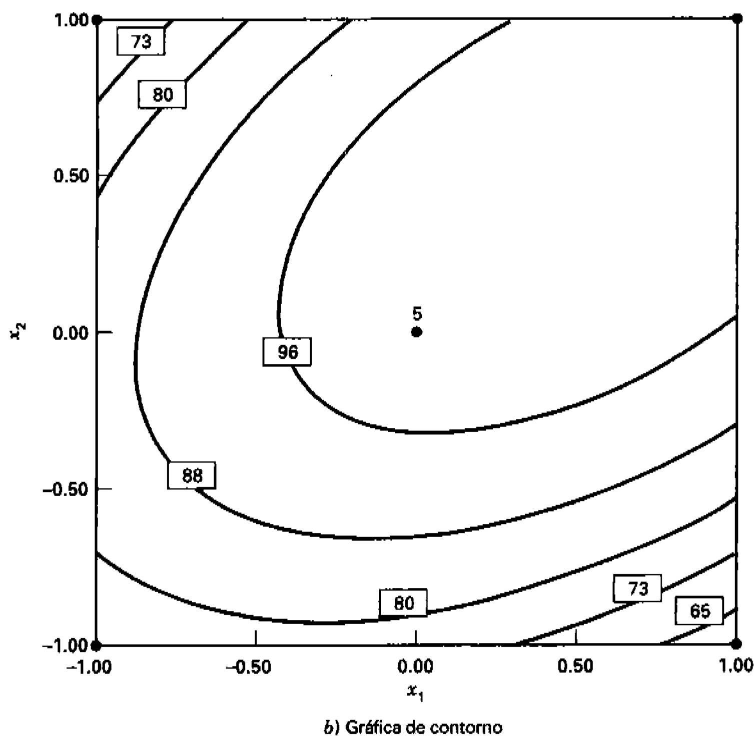
11-3.1 Localización del punto estacionario

Suponga que quieren encontrarse los niveles de $x_1, x_2, \ldots, x_k$ que optimizan la respuesta predicha. Este punto, en caso de existir, será el conjunto de las $x_1, x_2, \ldots, x_k$ para las que las derivadas parciales $\partial\hat{y}/\partial x_1 = \partial\hat{y}/\partial x_2 = \cdots = \partial\hat{y}/\partial x_k = 0$. A este punto, por ejemplo $x_{1,s}, x_{2,s}, \ldots, x_{k,s}$, se le llama **punto estacionario.** El punto estacionario podría representar 1) un punto de **respuesta máxima,** 2) un punto de **respuesta mínima,** o 3) un **punto silla.** Estas tres posibilidades se ilustran en las figuras 11-6 a 11-8.

Las gráficas de contorno desempeñan un papel muy importante en el estudio de las superficies de respuesta. Mediante la generación de gráficas de contorno utilizando software de computadora para el

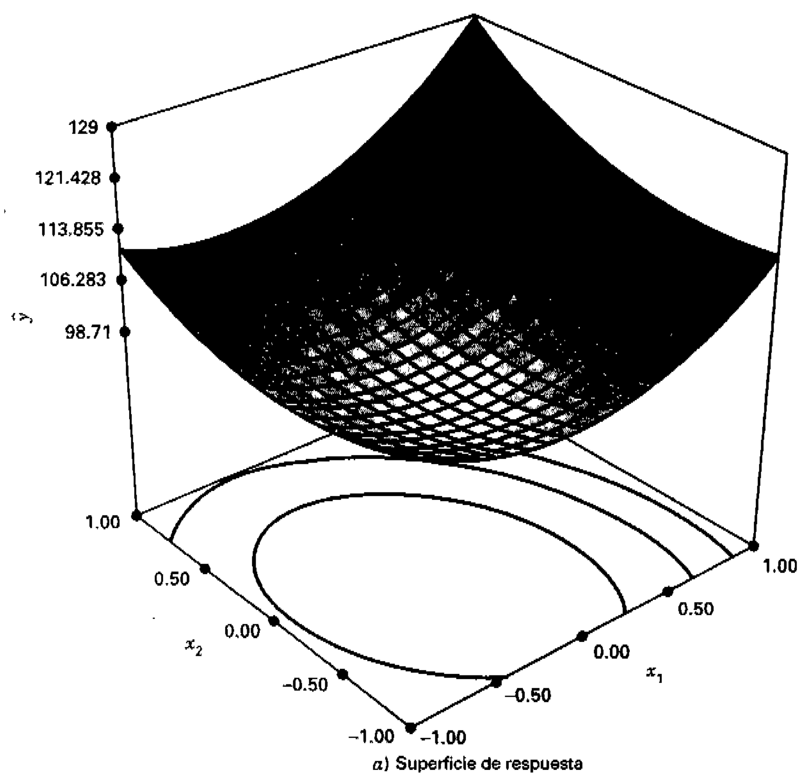


a) Superficie de respuesta

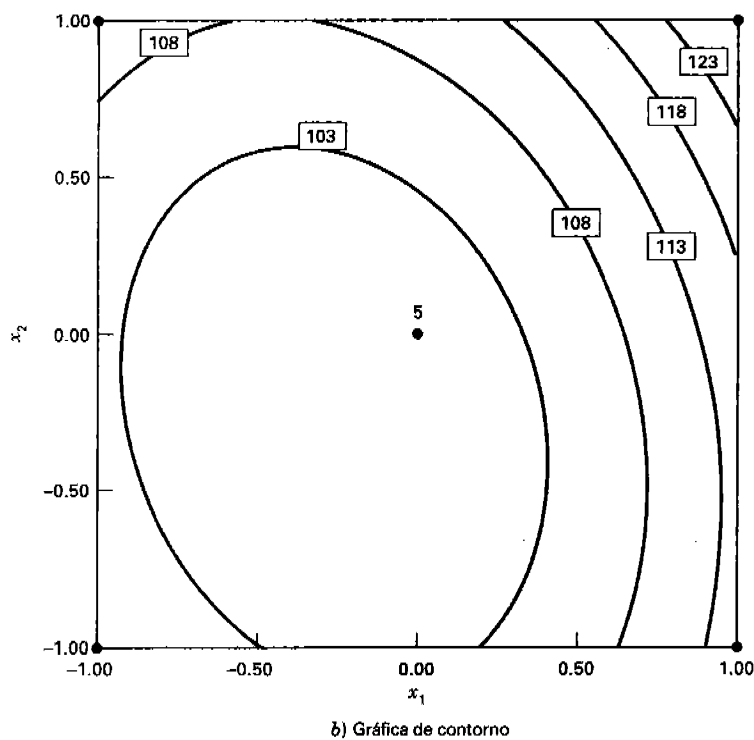


b) Gráfica de contorno

Figura 11-6 Superficie de respuesta y gráfica de contorno que ilustran una superficie con un máximo.

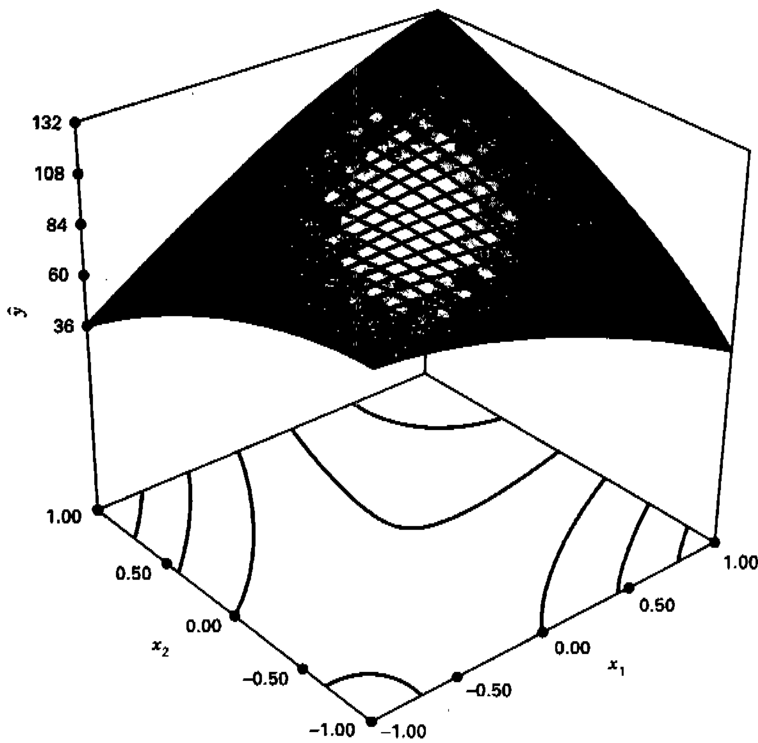


a) Superficie de respuesta

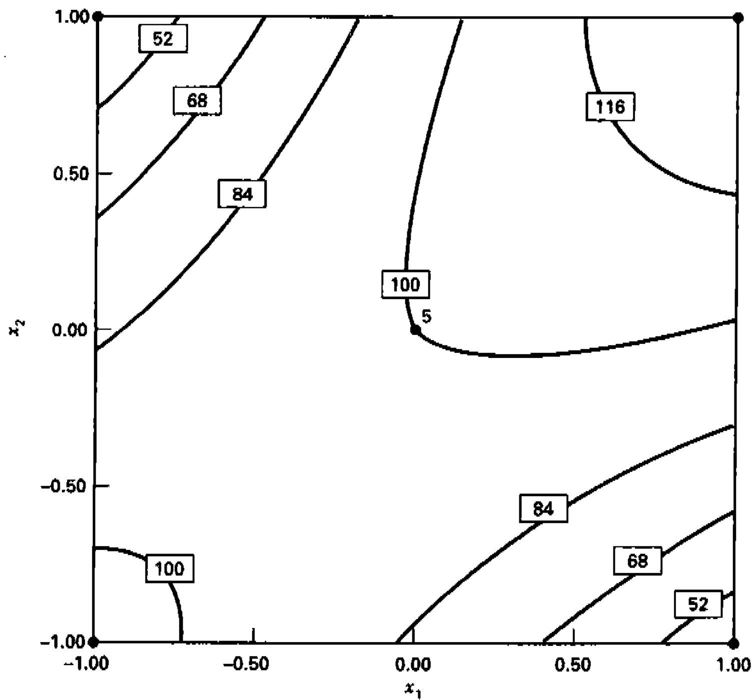


b) Gráfica de contorno

Figura 11-7 Superficie de respuesta y gráfica de contorno que ilustran una superficie con un mínimo.



a) Superficie de respuesta



b) Gráfica de contorno

Figura 11-8 Superficie de respuesta y gráfica de contorno que ilustran una superficie con un punto silla (o minimax).

análisis de superficie de respuesta, el experimentador puede por lo general caracterizar la forma de la superficie y localizar el óptimo con una precisión razonable.

Es posible obtener una solución matemática general para la localización del punto estacionario. Al escribir el modelo de segundo orden en notación matricial, se tiene

$$\hat{y} = \hat{\beta}_0 + \mathbf{x}'\mathbf{b} + \mathbf{x}'\mathbf{B}\mathbf{x} \tag{11-5}$$

donde

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix} \qquad \mathbf{b} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_k \end{bmatrix} \qquad \text{y} \qquad \mathbf{B} = \begin{bmatrix} \hat{\beta}_{11}, & \hat{\beta}_{12}/2, \ldots, & \hat{\beta}_{1k}/2 \\ & \hat{\beta}_{22}, \ldots, & \hat{\beta}_{2k}/2 \\ & \ddots & \\ \text{simétrica} & & \hat{\beta}_{kk} \end{bmatrix}$$

Es decir, **b** es un vector ($k \times 1$) de los coeficientes de regresión de primer orden y **B** es una matriz simétrica ($k \times k$) cuyos elementos de la diagonal principal son los coeficientes cuadráticos *puros* ($\hat{\beta}_{ii}$) y cuyos elementos que están fuera de la diagonal son la mitad de los coeficientes cuadráticos *mixtos* ($\hat{\beta}_{ij}$, $i \neq j$). La derivada de $\hat{y}$ con respecto a los elementos del vector **x** igualada con **0** es

$$\frac{\partial \hat{y}}{\partial \mathbf{x}} = \mathbf{b} + 2\mathbf{B}\mathbf{x} = \mathbf{0} \tag{11-6}$$

El punto estacionario es la solución de la ecuación 11-6, o

$$\mathbf{x}_s = -\tfrac{1}{2}\mathbf{B}^{-1}\mathbf{b} \tag{11-7}$$

Además, al sustituir la ecuación 11-7 en la 11-5, la respuesta predicha en el punto estacionario puede encontrarse como

$$\hat{y}_s = \hat{\beta}_0 + \tfrac{1}{2}\mathbf{x}_s'\mathbf{b} \tag{11-8}$$

11-3.2 Caracterización de la superficie de respuesta

Una vez que se ha encontrado el punto estacionario, generalmente es necesario caracterizar la superficie de respuesta en la vecindad inmediata de este punto. Por **caracterizar** se entiende determinar si el punto estacionario es el punto de una respuesta máxima, mínima o un punto silla. Por lo general también se desea estudiar la sensibilidad relativa de la respuesta a las variables $x_1, x_2, \ldots, x_k$.

Como ya se señaló, la forma más directa de hacer esto es examinando una **gráfica de contorno** del modelo ajustado. Si sólo hay dos o tres variables en el proceso (las x), la construcción e interpretación de esta gráfica de contorno es relativamente sencilla. Sin embargo, incluso cuando hay un número relativamente reducido de variables, un análisis más formal, llamado **análisis canónico,** puede ser útil.

Es conveniente transformar primero el modelo en un nuevo sistema de coordenadas con el origen en el punto estacionario **x**, y después hacer la rotación de los ejes de este sistema hasta que sean paralelos a los ejes principales de la superficie de respuesta ajustada. Esta transformación se ilustra en la figura 11-9. Puede demostrarse que se obtiene así el modelo ajustado

$$\hat{y} = \hat{y}_s + \lambda_1 w_1^2 + \lambda_2 w_2^2 + \cdots + \lambda_k w_k^2 \tag{11-9}$$

donde las $\{w_i\}$ son las variables independientes transformadas y las $\{\lambda_i\}$ son constantes. A la ecuación 11-9 se le llama la **forma canónica** del modelo. Además, las $\{\lambda_i\}$ son sólo **eigenvalores** o **raíces características** de la matriz **B**.

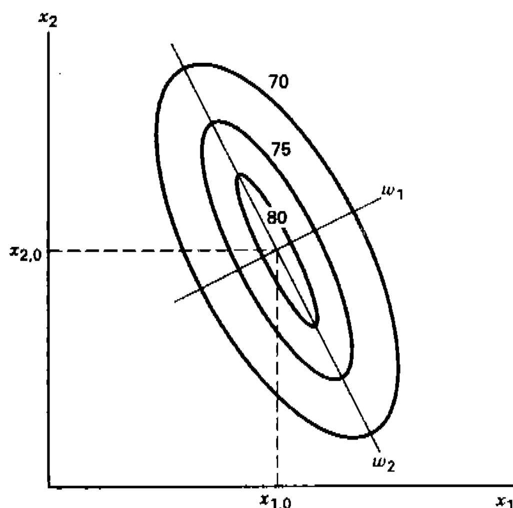


Figura 11-9 Forma canónica del modelo de segundo orden.

La naturaleza de la superficie de respuesta puede determinarse a partir del punto estacionario y de los *signos* y *magnitudes* de las $\{\lambda_i\}$. Primero suponga que el punto estacionario está dentro de la región de exploración para ajustar el modelo de segundo orden. Si todas las $\{\lambda_i\}$ son positivas, $\mathbf{x}_s$ es un punto de respuesta mínima; si todas las $\{\lambda_i\}$ son negativas, $\mathbf{x}_s$ es un punto de respuesta máxima; y si las $\{\lambda_i\}$ tienen signos diferentes, $\mathbf{x}_s$ es un punto silla. Además, la superficie presenta una inclinación mayor en la dirección w_i para la que $|\lambda_i|$ es el máximo. Por ejemplo, la figura 11-9 describe un sistema para el que $\mathbf{x}_s$ es un máximo (λ_1 y λ_2 son negativas) con $|\lambda_1| > |\lambda_2|$.

EJEMPLO 11-2 ··

Se continuará el análisis del proceso químico del ejemplo 11-1. No es posible ajustar un modelo de segundo orden en las variables x_1 y x_2 utilizando el diseño de la tabla 11-4. El experimentador decide aumentar este diseño con puntos suficientes para ajustar un modelo de segundo orden.[1] Obtiene cuatro observaciones en ($x_1 = 0, x_2 = \pm 1.414$) y ($x_1 = \pm 1.414, x_2 = 0$). El experimento completo se muestra en la tabla 11-6, y el diseño se ilustra en la figura 11-10. A este diseño se le llama **diseño central compuesto** (o DCC), el cual se estudiará con mayor detalle en la sección 11-4.2. En esta segunda fase del estudio, dos respuestas adicionales fueron de interés, la viscosidad y el peso molecular del producto. Las respuestas también se muestran en la tabla 11-6.

La atención se centrará en el ajuste de un modelo cuadrático para la respuesta rendimiento y_1 (las otras respuestas se analizarán en la sección 11-3.4). Por lo general se utiliza software de computadora para ajustar una superficie de respuesta y construir las gráficas de contorno. La tabla 11-7 contiene la salida de *Design-Expert*. Al examinar la tabla se observa que este paquete de software calcula primero las "sumas de cuadrados extra o secuenciales" de los términos lineales, cuadráticos y cúbicos del modelo (hay un mensaje de advertencia referente a los alias del modelo cúbico, ya que el DCC no contiene corridas suficientes para apoyar un modelo cúbico completo). Con base en el valor P pequeño de los términos cuadrá-

[1] El ingeniero corrió las cuatro observaciones adicionales aproximadamente en el mismo periodo en que corrió las nueve observaciones originales. Si hubiera transcurrido un lapso grande entre las dos series de corridas, habría sido necesaria la separación en bloques. La separación en bloques en los diseños de superficie de respuesta se revisa en la sección 11-4.3.

Tabla 11-6 Diseño central compuesto para el ejmplo 11-2

Variables naturales		Variables codificadas		Respuestas		
ξ_1	ξ_2	x_1	x_2	y_1 (rendimiento)	y_2 (viscosidad)	y_3 (peso molecular)
80	170	−1	−1	76.5	62	2940
80	180	−1	1	77.0	60	3470
90	170	1	−1	78.0	66	3680
90	180	1	1	79.5	59	3890
85	175	0	0	79.9	72	3480
85	175	0	0	80.3	69	3200
85	175	0	0	80.0	68	3410
85	175	0	0	79.7	70	3290
85	175	0	0	79.8	71	3500
92.07	175	1.414	0	78.4	68	3360
77.93	175	−1.414	0	75.6	71	3020
85	182.07	0	1.414	78.5	58	3630
85	167.93	0	−1.414	77.0	57	3150

ticos, se decide ajustar un modelo de segundo orden a la respuesta rendimiento. La salida de computadora muestra el modelo final en términos tanto de las variables codificadas como de los niveles naturales o reales de los factores.

En la figura 11-11 se muestra la gráfica de la superficie de respuesta tridimensional y la gráfica de contorno para la respuesta rendimiento en términos de las variables del proceso tiempo y temperatura. Es relativamente sencillo ver por el examen de estas figuras que el óptimo se encuentra muy cerca de 175°F y 85 minutos de tiempo de reacción y que la respuesta está en un máximo en este punto. Por el examen de la gráfica de contorno se observa que el proceso puede ser ligeramente más sensible a los cambios en el tiempo de reacción que a los cambios en la temperatura.

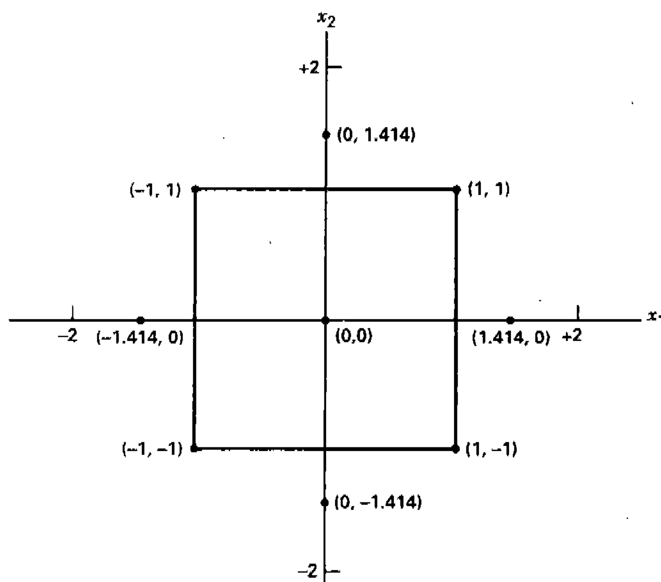


Figura 11-10 Diseño central compuesto para el ejemplo 11-2.

Tabla 11-7 Salida de computadora de *Design-Expert* para ajustar un modelo a la respuesta rendimiento del ejemplo 11-2

Response: yield
*****WARNING: The Cubic Model is Aliased!*****

Sequential Model Sum of Squares

Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Mean	80062.16	1	80062.16			
Linear	10.04	2	5.02	2.69	0.1166	
2FI	0.25	1	0.25	0.12	0.7350	
Quadratic	*17.95*	*2*	*8.98*	*126.88*	*<0.0001*	*Suggested*
Cubic	2.042E-003	2	1.021E-003	0.010	0.9897	Aliased
Residual	0.49	5	0.099			
Total	80090.90	13	6160.84			

"*Suma de cuadrados del modelo secuencial*": se selecciona el polinomio de orden más alto cuando los términos adicionales son significativos.

Lack of Fit Tests

Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Linear	18.49	6	3.08	58.14	0.0008	
2FI	18.24	5	3.65	68.82	0.0006	
Quadratic	*0.28*	*3*	*0.094*	*1.78*	*0.2897*	*Suggested*
Cubic	0.28	1	0.28	5.31	0.0826	Aliased
Pure Error	0.21	4	0.053			

"*Pruebas de falta de ajuste*": se quiere que el modelo seleccionado no tenga falta de ajuste significativa.

Model Summary Statistics

Source	Std. Dev.	R-Squared	Adjusted R-Squared	Predicted R-Squared	PRESS	
Linear	1.37	0.3494	0.2193	−0.0435	29.99	
2FI	1.43	0.3581	0.1441	−0.2730	36.59	
Quadratic	*0.27*	*0.9828*	*0.9705*	*0.9184*	*2.35*	*Suggested*
Cubic	0.31	0.9828	0.9588	0.3622	18.33	Aliased

"*Estadísticos de resumen del modelo*": se enfocan en el modelo que minimiza "PRESS" o, de manera equivalente, que maximiza la "R CUADRADA DE PREDICCIÓN".

Response: yield
ANOVA for Response Surface Quadratic Model
Analysis of variance table [Partial sum of squares]

Source	Sum of Squares	DF	Mean Square	F Value	Prob > F
Model	28.25	5	5.65	79.85	<0.0001
A	*7.92*	*1*	*7.92*	*111.93*	*<0.0001*
B	*2.12*	*1*	*2.12*	*30.01*	*0.0009*
A^2	*13.18*	*1*	*13.18*	*186.22*	*<0.0001*
B^2	*6.97*	*1*	*6.97*	*98.56*	*<0.0001*
AB	*0.25*	*1*	*0.25*	*3.53*	*0.1022*
Residual	0.50	7	0.071		
Lack of Fit	*0.28*	*3*	*0.094*	*1.78*	*0.2897*
Pure Error	*0.21*	*4*	*0.053*		
Cor Total	28.74	12			

Std. Dev.	0.27	R-Squared	0.9828
Mean	78.48	Adj R-Squared	0.9705
C.V.	0.34	Pred R-Squared	0.9184
PRESS	2.35	Adeq Precision	23.018

Tabla 11-7 (*continuación*)

Factor	Coefficient Estimate	DF	Standard Error	95% CI Low	95% CI High	VIF
Intercept	79.94	1	0.12	79.66	80.22	
A-time	0.99	1	0.094	0.77	1.22	1.00
B-temp	0.52	1	0.094	0.29	0.74	1.00
A^2	−1.38	1	0.10	−1.61	−1.14	1.02
B^2	−1.00	1	0.10	−1.24	−0.76	1.02
AB	0.25	1	0.13	−0.064	0.56	1.00

Final Equation in Terms of Coded Factors:

yield =
+79.94
+0.99 * A
+0.52 * B
−1.38 * A^2
−1.00 * B^2
+0.25 * A * B

Final Equation in Terms of Actual Factors:

yield =
−1430.52285
+7.80749 * time
+13.27053 * temp
−0.055050 * time^2
−0.040050 * temp^2
+0.010000 * time * temp

Diagnostics Case Statistics

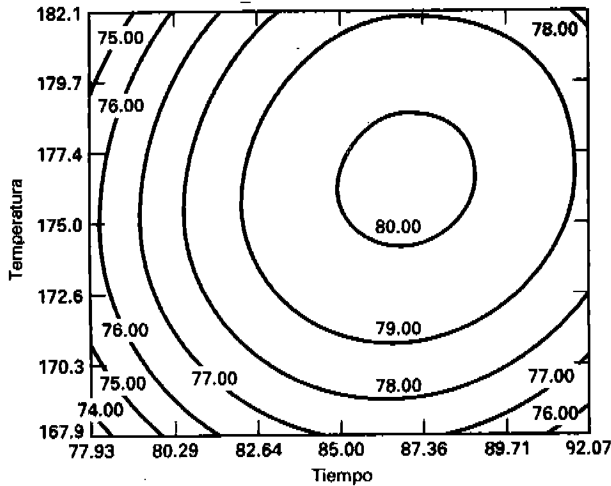
Run Order	Standard Order	Actual Value	Predicted Value	Residual	Leverage	Student Residual	Cook's Distance	Outlier t
8	1	76.50	76.30	0.20	0.625	1.213	0.409	1.264
6	2	78.00	77.79	0.21	0.625	1.275	0.452	1.347
9	3	77.00	76.83	0.17	0.625	1.027	0.293	1.032
11	4	79.50	79.32	0.18	0.625	1.089	0.329	1.106
12	5	75.60	75.78	−0.18	0.625	−1.107	0.341	−1.129
10	6	78.40	78.59	−0.19	0.625	−1.195	0.396	−1.240
7	7	77.00	77.21	−0.21	0.625	−1.283	0.457	−1.358
1	8	78.50	78.67	−0.17	0.625	−1.019	0.289	−1.023
5	9	79.90	79.94	−0.040	0.200	−0.168	0.001	−0.156
3	10	80.30	79.94	0.36	0.200	1.513	0.095	1.708
13	11	80.00	79.94	0.060	0.200	0.252	0.003	0.235
2	12	79.70	79.94	−0.24	0.200	−1.009	0.042	−1.010
4	13	79.80	79.94	−0.14	0.200	−0.588	0.014	−0.559

La localización del punto estacionario también podría encontrarse utilizando la solución general de la ecuación 11-7. Observe que

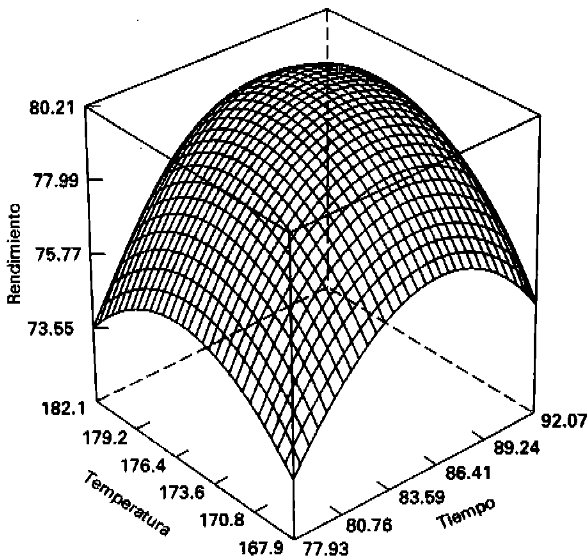
$$\mathbf{b} = \begin{bmatrix} 0.995 \\ 0.515 \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} -1.376 & 0.1250 \\ 0.1250 & -1.001 \end{bmatrix}$$

y, por la ecuación 11-7, el punto estacionario es

$$\begin{aligned} \mathbf{x}_s &= -\tfrac{1}{2}\mathbf{B}^{-1}\mathbf{b} \\ &= -\tfrac{1}{2}\begin{bmatrix} -0.7345 & -0.0917 \\ -0.0917 & -1.0096 \end{bmatrix}\begin{bmatrix} 0.995 \\ 0.515 \end{bmatrix} = \begin{bmatrix} 0.389 \\ 0.306 \end{bmatrix} \end{aligned}$$



a) La gráfica de contorno



b) La gráfica de superficie de respuesta

Figura 11-11 Gráficas de contorno y de superficie de respuesta de la respuesta rendimiento, ejemplo 11-2.

Es decir, $x_{1,s} = 0.389$ y $x_{2,s} = 0.306$. En términos de las variables naturales, el punto estacionario es

$$0.389 = \frac{\xi_1 - 85}{5} \qquad 0.306 = \frac{\xi_2 - 175}{5}$$

de donde se obtiene $\xi_1 = 86.95 \simeq 87$ minutos de tiempo de reacción y $\xi_2 = 176.53 \simeq 176.5$°F. Este valor está muy cerca del punto estacionario que se encontró por examen visual en la gráfica de contorno de la figura 11-11. Al utilizar la ecuación 11-8, la respuesta predicha en el punto estacionario puede encontrarse como $\hat{y}_s = 80.21$.

El análisis canónico que se describe en esta sección también puede usarse para caracterizar la superficie de respuesta. Primero es necesario expresar el modelo ajustado en la forma canónica (ecuación 11-9). Los eigenvalores λ_1 y λ_2 son las raíces de la ecuación de determinantes

$$|\mathbf{B}-\lambda\mathbf{I}| = 0$$

$$\begin{vmatrix} -1.376-\lambda & 0.1250 \\ 0.1250 & -1.001-\lambda \end{vmatrix} = 0$$

que se reduce a

$$\lambda^2 + 2.3788\lambda + 1.3639 = 0$$

Las raíces de esta ecuación cuadrática son $\lambda_1 = -0.9641$ y $\lambda_2 = -1.4147$. Por lo tanto, la forma canónica del modelo ajustado es

$$\hat{y} = 80.21 - 0.9641w_1^2 - 1.4147w_2^2$$

Puesto que tanto λ_1 como λ_2 son negativas y el punto estacionario está en la región de exploración, se concluye que el punto estacionario es un máximo.

....

En algunos problemas MSR puede ser necesario encontrar la relación entre las **variables canónicas** $\{w_i\}$ y las **variables del diseño** $\{x_i\}$. Esto es particularmente cierto cuando es imposible operar el proceso en el punto estacionario. Como una ilustración, suponga que en el ejemplo 11-2 el proceso no pudo operarse en $\xi_1 = 87$ minutos y $\xi_2 = 176.5$°F debido a que esta combinación de factores resulta en un costo excesivo. Se quiere "regresar" ahora del punto estacionario a un punto con un costo menor sin incurrir en pérdidas considerables en el rendimiento. La forma canónica del modelo indica que la superficie es menos sensible a la pérdida de rendimiento en la dirección w_1. La exploración de la forma canónica requiere convertir los puntos del espacio (w_1, w_2) en puntos del espacio (x_1, x_2).

En general, las variables $\mathbf{x}$ se relacionan con las variables canónicas $\mathbf{w}$ por

$$\mathbf{w} = \mathbf{M}'(\mathbf{x} - \mathbf{x}_s)$$

donde $\mathbf{M}$ es una matriz ortogonal $(k \times k)$. Las columnas de $\mathbf{M}$ son los eigenvectores normalizados asociados con $\{\lambda_i\}$. Es decir, si $\mathbf{m}_i$ es la columna i-ésima de $\mathbf{M}$, entonces $\mathbf{m}_i$ es la solución de

$$(\mathbf{B} - \lambda_i\mathbf{I})\mathbf{m}_i = \mathbf{0} \qquad (11\text{-}10)$$

para la que $\Sigma_{j=1}^{k} m_{ji}^2 = 1$.

El procedimiento se ilustra usando el modelo de segundo orden ajustado del ejemplo 11-2. Para $\lambda_1 = -0.9641$, la ecuación 11-10 queda como

$$\begin{bmatrix} (-1.376+0.9641) & 0.1250 \\ 0.1250 & (-1.001+0.9641) \end{bmatrix} \begin{bmatrix} m_{11} \\ m_{21} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

o

$$-0.4129m_{11} + 0.1250m_{21} = 0$$
$$0.1250m_{11} - 0.0377m_{21} = 0$$

Quiere obtenerse la solución normalizada de estas ecuaciones, es decir, aquella para la que $m_{11}^2 + m_{21}^2 = 1$. No existe una solución única para estas ecuaciones, por lo que lo más conveniente es asignar un valor ar-

bitrario a una de las incógnitas, resolver el sistema y normalizar la solución. Al hacer $m_{21}^* = 1$, se encuentra $m_{11}^* = 0.3027$. Para normalizar esta solución, m_{11}^* y m_{21}^* se dividen entre

$$\sqrt{(m_{11}^*)^2 + (m_{21}^*)^2} = \sqrt{(0.3027)^2 + (1)^2} = 1.0448$$

Se obtiene así la solución normalizada

$$m_{11} = \frac{m_{11}^*}{1.0448} = \frac{0.3027}{1.0448} = 0.2897$$

y

$$m_{21} = \frac{m_{21}^*}{1.0448} = \frac{1}{1.0448} = 0.9571$$

que es la primera columna de la matriz **M**.

Utilizando $\lambda_2 = -1.4147$ puede repetirse el procedimiento anterior, obteniéndose $m_{12} = -0.9574$ y $m_{22} = 0.2888$ como la segunda columna de **M**. Por lo tanto, se tiene

$$\mathbf{M} = \begin{bmatrix} 0.2897 & -0.9574 \\ 0.9571 & 0.2888 \end{bmatrix}$$

La relación entre las variables **w** y **x** es

$$\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} 0.2897 & 0.9571 \\ -0.9574 & 0.2888 \end{bmatrix} \begin{bmatrix} x_1 - 0.389 \\ x_2 - 0.306 \end{bmatrix}$$

o

$$w_1 = 0.2897(x_1 - 0.389) + 0.9571(x_2 - 0.306)$$
$$w_2 = -0.9574(x_1 - 0.389) + 0.2888(x_2 - 0.306)$$

Si quisiera explorarse la superficie de respuesta en la vecindad del punto estacionario, podrían determinarse los puntos apropiados en los cuales hacer las observaciones en el espacio (w_1, w_2) y usar después la relación anterior para convertir estos puntos en el espacio (x_1, x_2) para que puedan realizarse las corridas.

11-3.3 Sistemas de cordilleras

No es raro encontrar variaciones de las superficies de respuesta con máximos o mínimos puros o con puntos silla estudiadas en la sección anterior. Los sistemas de cordilleras, en particular, son muy comunes. Considere la forma canónica del modelo de segundo orden presentado anteriormente en la ecuación 11-9:

$$\hat{y} = \hat{y}_s + \lambda_1 w_1^2 + \lambda_2 w_2^2 + \cdots + \lambda_k w_k^2$$

Suponga ahora que el punto estacionario $\mathbf{x}_s$ está dentro de la región de experimentación; además, sea que una o más de las λ_i sean muy pequeñas (por ejemplo, $\lambda_i \simeq 0$). Entonces la variable de respuesta es muy insensible a las variables w_i multiplicadas por las λ_i pequeñas.

En la figura 11-12 se presenta una gráfica de contorno en la que se ilustra esta situación para $k = 2$ variables con $\lambda_1 = 0$. (En la práctica, λ_1 estaría cerca de cero pero no sería exactamente igual a cero.) En teoría, el modelo canónico para esta superficie de respuesta es

$$\hat{y} = \hat{y}_s + \lambda_2 w_2^2$$

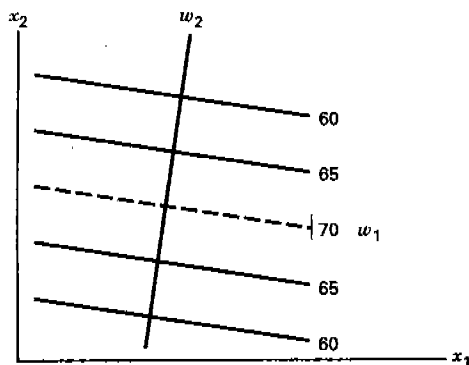


Figura 11-12 Gráfica de contorno de un sistema de cordilleras estacionarias.

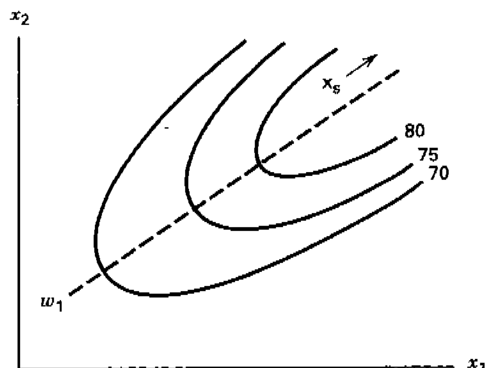


Figura 11-13 Gráfica de contorno de un sistema de cordilleras crecientes.

con λ_2 negativa. Observe que el marcado estiramiento en la dirección w_1 ha resultado en una línea de centros en $\hat{y} = 70$ y el óptimo puede tomarse en cualquier lugar a lo largo de esta línea. A este tipo de superficie de respuesta se le llama **sistema de cordilleras estacionarias.**

Si el punto estacionario está muy apartado de la región de exploración para el ajuste del modelo de segundo orden y una λ_i (o más) está cerca de cero, entonces la superficie puede ser un **sistema de cordilleras crecientes.** En la figura 11-13 se ilustra una cordillera creciente para $k = 2$ variables con λ_1 cerca de cero y λ_2 negativa. En este tipo de sistema de cordilleras no pueden hacerse inferencias acerca de la verdadera superficie o del punto estacionario porque $\mathbf{x}_s$ está fuera de la región donde se ha ajustado el modelo. Sin embargo, la exploración adicional está garantizada en la dirección w_1. Si λ_2 hubiera sido positiva, este sistema se habría llamado cordillera descendente.

11-3.4 Respuestas múltiples

Muchos problemas de superficies de respuesta incluyen el análisis de varias respuestas, como en el ejemplo 11-2, donde el experimentador midió tres. En dicho ejemplo, el proceso se optimizó únicamente con respecto a la respuesta rendimiento y_1.

La consideración simultánea de respuestas múltiples requiere construir primero un modelo de superficie de respuesta apropiado para cada respuesta y después intentar encontrar un conjunto de condiciones de operación que optimice en cierto sentido todas las respuestas o que al menos las mantenga en los rangos deseados. Un estudio completo del problema de las respuestas múltiples se ofrece en Myers y Montgomery [85a].

En el ejemplo 11-2 pueden obtenerse modelos para las respuestas viscosidad y peso molecular (y_2 y y_3, respectivamente) de la siguiente manera:

$$\hat{y}_2 = 70.00 - 0.16x_1 - 0.95x_2 - 0.69x_1^2 - 6.69x_2^2 - 1.25x_1x_2$$
$$\hat{y}_3 = 3386.2 + 205.1x_1 + 17.4x_2$$

En términos de los niveles naturales del tiempo (ξ_1) y la temperatura (ξ_2), estos modelos son

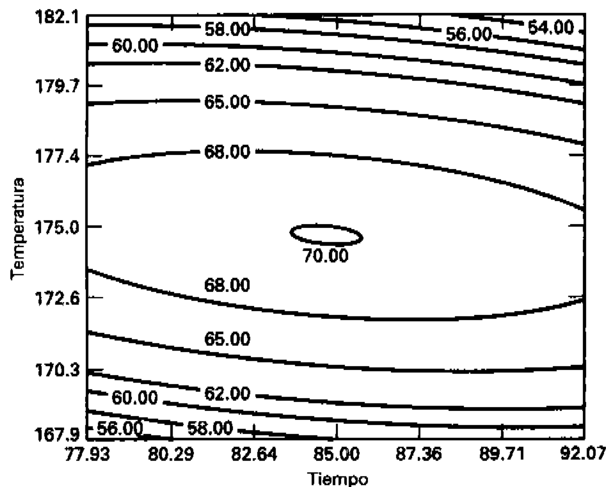
$$\hat{y}_2 = -9030.74 + 13.393\xi_1 + 97.708\xi_2$$
$$- 2.75\times10^{-2}\xi_1^2 - 0.26757\xi_2^2 - 5\times10^{-2}\xi_1\xi_2$$

y

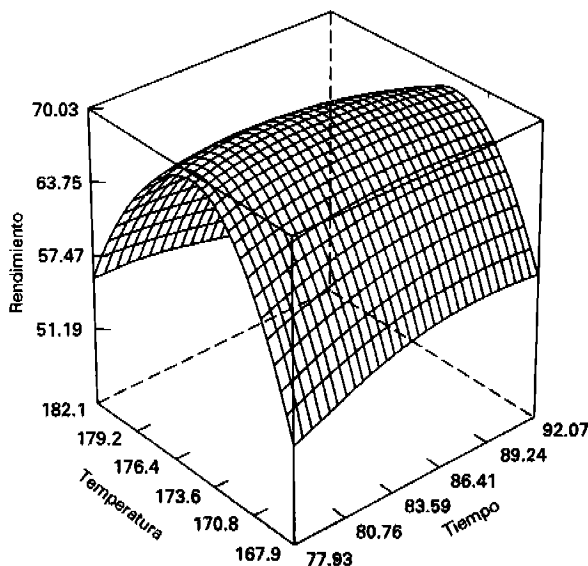
$$\hat{y}_3 = -6308.8 + 41.025\xi_1 + 35.473\xi_2$$

En las figuras 11-14 y 11-15 se presentan las gráficas de contorno y superficie de respuesta para estos modelos.

Un enfoque relativamente directo para optimizar varias respuestas que funciona bien cuando sólo hay pocas variables en el proceso es la **superposición de las gráficas de contorno** de cada respuesta. En la figura 11-16 se muestra una gráfica de superposición para las tres respuestas del ejemplo 11-2, con los contornos para los que y_1 (rendimiento) ≥ 78.5, $62 \leq y_2$ (viscosidad) ≤ 68, y y_3 (peso molecular Mn) ≤ 3400. Si estos límites representan condiciones importantes que el proceso debe satisfacer, entonces, como se muestra en la porción no sombreada de la figura 11-16, existen varias combinaciones del tiempo y la temperatura que resultarán en un proceso satisfactorio. El experimentador puede hacer el examen visual de

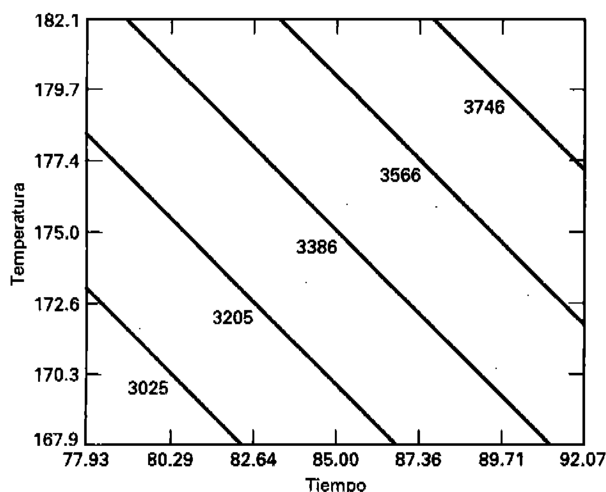


a) La gráfica de contorno

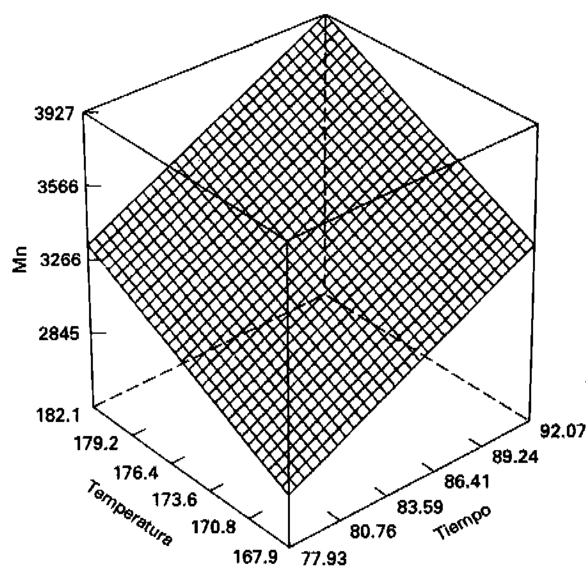


b) La gráfica de la superficie de respuesta

Figura 11-14 Gráfica de contorno y gráfica de la superficie de respuesta de la viscosidad, ejemplo 11-2.



a) La gráfica de contorno



b) La gráfica de la superficie de respuesta

Figura 11-15 Gráfica de contorno y gráfica de la superficie de respuesta del peso molecular, ejemplo 11-2.

la gráfica de contorno para determinar las condiciones de operación apropiadas. Por ejemplo, es posible que el experimentador esté más interesado en la región más grande de las dos regiones factibles que se muestran en la figura 11-16.

Cuando hay más de tres variables del diseño, se hace muy complicada la superposición de las gráficas de contorno, ya que la gráfica de contorno es bidimensional, y $k-2$ de las variables del diseño deben mantenerse constantes para construir la gráfica. Con frecuencia se necesita una gran cantidad de ensayo y error para determinar cuáles son los factores que deben mantenerse constantes y qué niveles seleccionar para obtener la mejor vista de la superficie. Por lo tanto, existe interés práctico en métodos de optimización más formales para las respuestas múltiples.

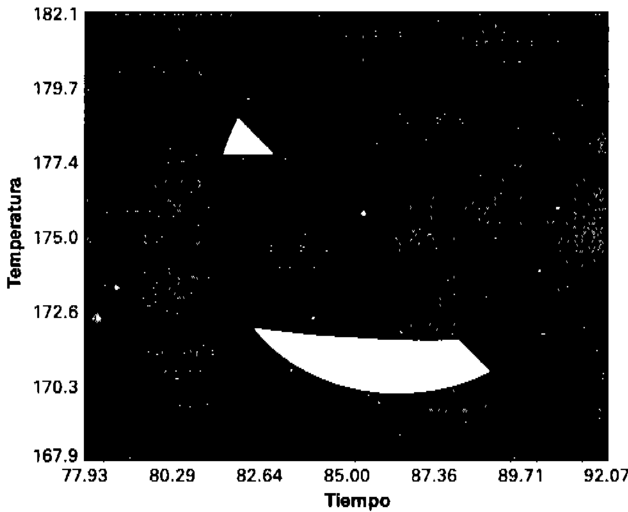


Figura 11-16 Región del óptimo encontrada superponiendo las superficies de respuesta del rendimiento, la viscosidad y el peso molecular, ejemplo 11-2.

Un enfoque popular consiste en formular y resolver el problema como un **problema de optimización restringida.** Para ilustrar este enfoque utilizando el ejemplo 11-2, el problema podría formularse como

$$\text{Máx } y_1$$
$$\text{sujeto a}$$
$$62 \leq y_2 \leq 68$$
$$y_3 \leq 3400$$

Se cuenta con varias técnicas numéricas que pueden usarse para resolver este problema. En ocasiones se hace referencia a estas técnicas como **métodos de programación no lineal.** El paquete de software *Design-Expert* resuelve esta versión del problema utilizando un procedimiento de búsqueda directa. Las dos soluciones encontradas son

$$\text{tiempo} = 83.5 \qquad \text{temperatura} = 177.1 \qquad \hat{y}_1 = 79.5$$

y

$$\text{tiempo} = 86.6 \qquad \text{temperatura} = 172.25 \qquad \hat{y}_1 = 79.5$$

Observe que la primera solución es la región factible superior (la más pequeña) del espacio del diseño (referirse a la figura 11-16), mientras que la segunda solución es la región más grande. Ambas soluciones están muy cerca de los límites de las restricciones.

Otro enfoque útil para la optimización de respuestas múltiples es usar la técnica de optimización simultánea popularizada por Derringer y Suich [37]. Su procedimiento hace uso de las **funciones con condición de deseable.** El enfoque general consiste en convertir primero cada respuesta y_i en una función con condición de deseable individual d_i que varía en el rango

$$0 \leq d_i \leq 1$$

donde si la respuesta y_i está en su meta u objetivo, entonces $d_i = 1$, y si la respuesta está fuera de una región aceptable, $d_i = 0$. Después las variables del diseño se eligen para maximizar la condición de deseable global

$$D = (d_1 \cdot d_2 \cdot \dots \cdot d_m)^{1/m}$$

donde hay m respuestas.

Las funciones con condición de deseable individual están estructuradas como se indica en la figura 11-17. Si el objetivo T para la respuesta y es un valor máximo,

$$d = \begin{cases} 0 & y < L \\ \left(\dfrac{y-L}{T-L}\right)^r & L \le y \le T \\ 1, & y > T \end{cases} \tag{11-11}$$

cuando la ponderación $r = 1$, la función con condición de deseable es lineal. Al elegir $r > 1$ se pone más interés en estar cerca del valor objetivo, y cuando se elige $0 < r < 1$ esto tiene menos importancia. Si el objetivo para la respuesta es un valor mínimo,

$$d = \begin{cases} 1 & y < T \\ \left(\dfrac{U-y}{U-T}\right)^r & T \le y \le U \\ 0 & y > U \end{cases} \tag{11-12}$$

La función con condición de deseable de dos colas que se muestra en la figura 11-17*c* supone que el objetivo se localiza entre los límites inferior (L) y superior (U), y se define como

$$d = \begin{cases} 0 & y < L \\ \left(\dfrac{y-L}{T-L}\right)^{r_1} & L \le y \le T \\ \left(\dfrac{U-y}{U-T}\right)^{r_2} & T \le y \le U \\ 0 & y > U \end{cases} \tag{11-13}$$

Se usó el paquete de software *Design-Expert* para resolver el ejemplo 11-2 utilizando el enfoque de la función con condición de deseable. Se eligió $T = 80$ como el objetivo para la respuesta rendimiento, $U = 70$, y se fijó la ponderación de esta condición de deseable individual igual a la unidad. Se hizo $T = 65$ para la respuesta viscosidad con $L = 62$ y $U = 68$ (para ser consistente con las especificaciones), con ambas ponderaciones $r_1 = r_2 = 1$. Por último, se indicó que cualquier peso molecular abajo de 3400 era aceptable. Se encontraron dos soluciones.

Solución 1:

$$\text{Tiempo} = 86.5 \quad \text{Temperatura} = 170.5 \quad D = 0.822$$
$$\hat{y}_1 = 78.8 \quad \hat{y}_2 = 65 \quad \hat{y}_3 = 3287$$

Solución 2:

$$\text{Tiempo} = 82 \quad \text{Temperatura} = 178.8 \quad D = 0.792$$
$$\hat{y}_1 = 78.5 \quad \hat{y}_2 = 65 \quad \hat{y}_3 = 3400$$

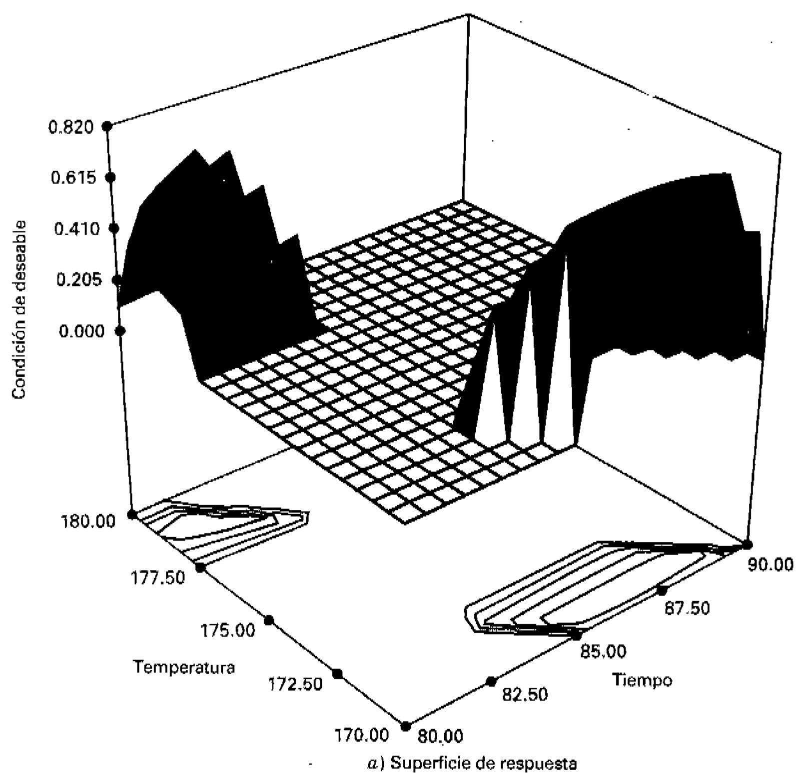
La solución 1 tiene la condición de deseable global más alta. Observe que resulta en una viscosidad acorde con el objetivo y en un peso molecular aceptable. Esta solución está contenida en la más grande de las dos regiones de operación de la figura 11-16, mientras que la segunda solución está contenida en la región más pequeña. En la figura 11-18 se muestran las gráficas de la superficie de respuesta y de contorno de la función con condición de deseable global D.

a) El objetivo (blanco) es maximizar *y*

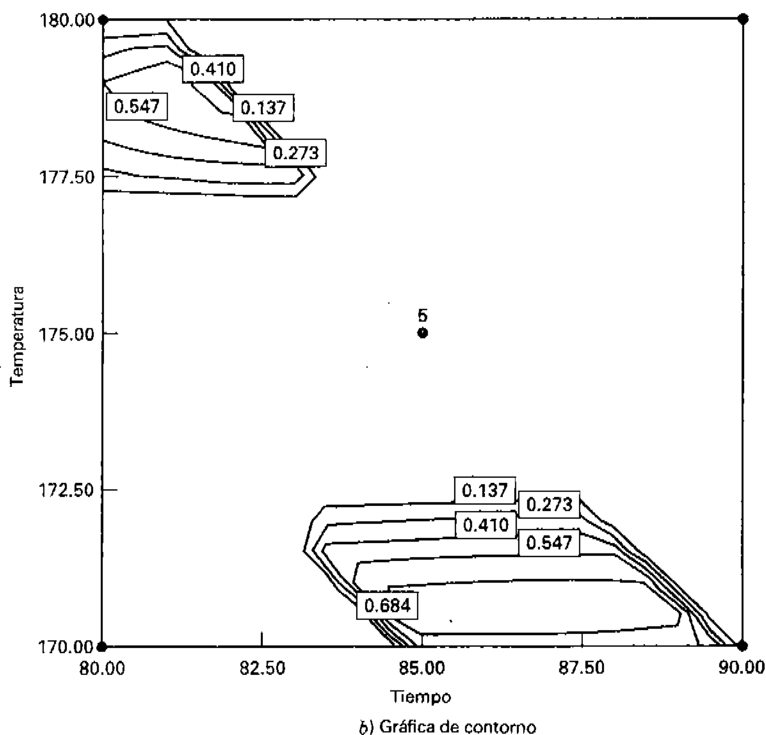
b) El objetivo (blanco) es minimizar *y*

c) El objetivo (blanco) es que *y* esté tan cerca como sea posible de la especificación

Figura 11-17 Funciones con condición de deseables individuales para la optimización simultánea.



a) Superficie de respuesta



b) Gráfica de contorno

Figura 11-18 Gráfica de la superficie de respuesta y de contorno de la función con condición de deseable del ejemplo 11-2.

11-4 DISEÑOS EXPERIMENTALES PARA AJUSTAR SUPERFICIES DE RESPUESTA

El ajuste y análisis de superficies de respuesta se facilita en gran medida con la elección apropiada del diseño experimental. En esta sección se revisan algunos aspectos de la selección del diseño apropiado para ajustar superficies de respuesta.

Cuando se selecciona un diseño de superficie de respuesta, algunas de las características deseables en el diseño son las siguientes:

1. Proporciona una distribución razonable de los puntos de los datos (y en consecuencia información) en toda la región de interés.
2. Permite que se investigue la adecuación del modelo, incluyendo la falta de ajuste.
3. Permite que los experimentos se realicen en bloques.
4. Permite que los diseños de orden superior se construyan secuencialmente.
5. Proporciona una estimación interna del error.
6. Proporciona estimaciones precisas de los coeficientes del modelo.
7. Proporciona un buen perfil de la varianza de predicción en toda la región experimental.
8. Proporciona una robustez razonable contra los puntos atípicos o los valores faltantes.
9. No requiere un gran número de corridas.
10. No requiere demasiados niveles de las variables independientes.
11. Asegura la simplicidad del cálculo de los parámetros del modelo.

Estas características entran en conflicto en ocasiones, por lo que con frecuencia debe aplicarse la discrecionalidad al seleccionar un diseño. Para mayor información sobre la elección de un diseño de superficie de respuesta, referirse a Myers y Montgomery [85a], Box y Draper [16b] y Khuri y Cornell [67].

11-4.1 Diseños para ajustar el modelo de primer orden

Suponga que quiere ajustarse el modelo de primer orden en k variables

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \varepsilon \tag{11-14}$$

Hay una clase única de diseños que minimizan la varianza de los coeficientes de regresión $\{\hat{\beta}_i\}$. Se trata de los **diseños de primer orden ortogonales.** Un diseño de primer orden es ortogonal si todos los elementos que están fuera de la diagonal de la matriz $(\mathbf{X}'\mathbf{X})$ son cero. Esto implica que la suma de los productos cruzados de las columnas de la matriz $\mathbf{X}$ sea cero.

La clase de los diseños de primer orden ortogonales incluye los factoriales 2^k y las fracciones de la serie 2^k en las que los efectos principales no son alias entre sí. Al usar estos diseños se supone que los niveles bajo y alto de los k factores están codificados en los niveles usuales ± 1.

El diseño 2^k no permite la estimación del error experimental a menos que se hagan réplicas de algunas corridas. Un método común de incluir las réplicas en el diseño 2^k es aumentar el diseño con varias observaciones en el centro (el punto $x_i = 0$, $i = 1, 2, ..., k$). La adición de puntos centrales al diseño 2^k no influye en las $\{\hat{\beta}_i\}$ para $i \geq 1$, pero la estimación de β_0 se convierte en el gran promedio de todas las observaciones. Además, la adición de puntos centrales no altera la propiedad de ortogonalidad del diseño. En

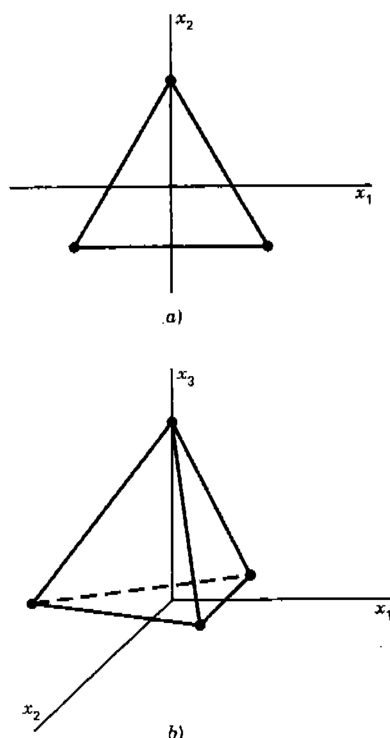


Figura 11-19 El diseño símplex para *a*) $k = 2$ variables y *b*) $k = 3$ variables.

el ejemplo 11-1 se ilustra el uso de un diseño 2^2 aumentado con cinco puntos centrales para ajustar un modelo de primer orden.

Otro diseño de primer orden ortogonal es el **diseño símplex.** El diseño símplex es una figura de lados regulares con $k + 1$ vértices en k dimensiones. Por lo tanto, el diseño símplex para $k = 2$ es un triángulo equilátero, y para $k = 3$ es un tetraedro regular. En la figura 11-19 se muestran diseños símplex de dos y tres dimensiones.

11-4.2 Diseños para ajustar el modelo de segundo orden

En el ejemplo 11-2 se hizo la introducción informal (e incluso antes en el ejemplo 6-6) del **diseño central compuesto** o **DCC** para ajustar un modelo de segundo orden. Se trata de la clase más popular de diseños usados para ajustar estos modelos. En general, el DCC consta de un factorial 2^k (o de un factorial fraccionado de resolución V) con n_F corridas, $2k$ corridas axiales o estrella y n_C corridas centrales. En la figura 11-20 se muestra el DCC para $k = 2$ y $k = 3$ factores.

El despliegue práctico de un DCC surge con frecuencia a través de la **experimentación secuencial,** como en los ejemplos 11-1 y 11-2. Es decir, se ha usado un diseño 2^k para ajustar un modelo de primer orden, este modelo ha presentado falta de ajuste, y después se agregaron las corridas axiales para permitir la incorporación de los términos cuadráticos en el modelo. El DCC es un diseño muy eficiente para ajustar el modelo de segundo orden. Hay dos parámetros en el diseño que deben especificarse: la distancia α de las corridas axiales al centro del diseño y el número de puntos centrales n_C. A continuación se analiza la elección de estos dos parámetros.

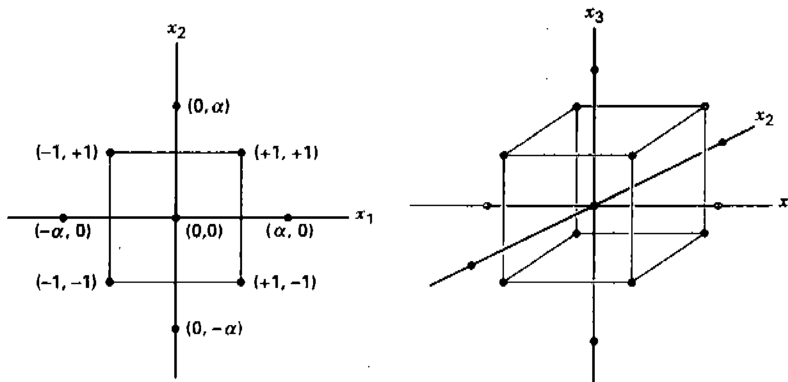


Figura 11-20 Diseños centrales compuestos para $k = 2$ y $k = 3$.

Rotabilidad

Es importante que el modelo de segundo orden proporcione buenas predicciones en toda la región de interés. Una manera de definir "buenas" es requerir que el modelo tenga una varianza razonablemente consistente y estable de la respuesta predicha en los puntos de interés **x**. Recuerde, por la ecuación 10-40, que la varianza de la respuesta predicha en algún punto **x** es

$$V[\hat{y}(\mathbf{x})] = \sigma^2 \mathbf{x}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}$$

Box y Hunter [17a] propusieron que un diseño de superficie de respuesta de segundo orden debe ser **rotable.** Esto significa que la $V[\hat{y}(\mathbf{x})]$ es la misma en todos los puntos **x** que están a la misma distancia del centro del diseño. Es decir, la varianza de la respuesta predicha es constante en esferas.

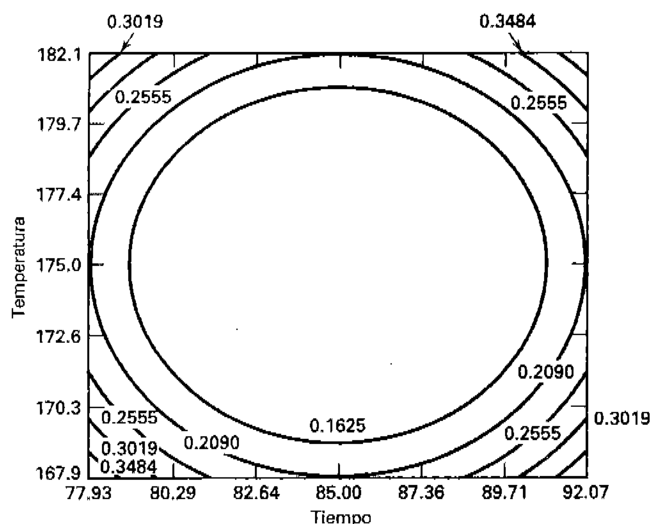
En la figura 11-21 se muestran los contornos de $\sqrt{V[\hat{y}(\mathbf{x})]}$ constante para el ajuste del modelo de segundo orden utilizando el DCC en el ejemplo 11-2. Observe que los contornos de desviación estándar constante de la respuesta predicha son círculos concéntricos. Un diseño con esta propiedad dejará la varianza de $\hat{y}$ sin cambio cuando el diseño se rota alrededor del centro (0, 0, ..., 0), de ahí el nombre de diseño *rotable*.

La rotabilidad es una base razonable para la selección de un diseño de superficie de respuesta. Puesto que la finalidad de la MSR es la optimización, y la localización del óptimo se desconoce antes de correr el experimento, tiene sentido el uso de un diseño que proporcione una precisión de estimación igual en todas las direcciones (puede demostrarse que cualquier diseño de primer orden ortogonal es rotable).

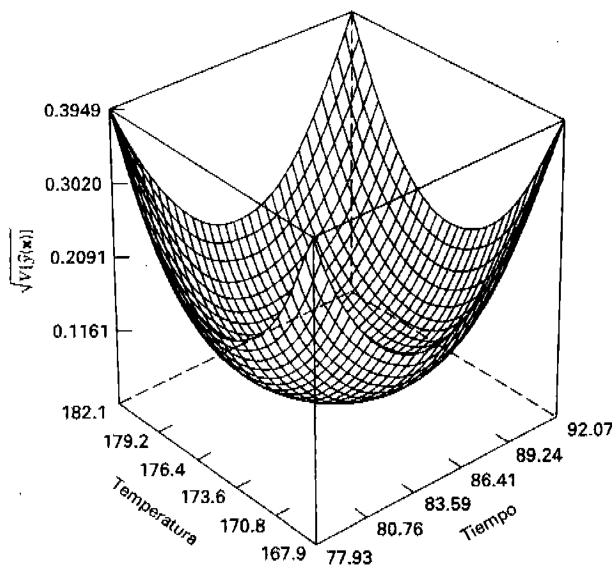
Un diseño central compuesto se hace rotable mediante la elección de α. El valor de α para la rotabilidad depende del número de puntos en la porción factorial del diseño; de hecho, $\alpha = (n_F)^{1/4}$ produce un diseño central compuesto rotable, donde n_F es el número de puntos usados en la porción factorial del diseño.

El DCC esférico

La rotabilidad es una propiedad esférica; es decir, tiene mayor sentido como criterio de diseño cuando la región de interés es una esfera. Sin embargo, no es importante tener una rotabilidad exacta para tener un buen diseño. De hecho, para una región esférica de interés, la mejor elección de α desde el punto de vista de la varianza de predicción para el DCC es hacer $\alpha = \sqrt{k}$. Este diseño, llamado **DCC esférico,** coloca todos los puntos factoriales y axiales del diseño sobre la superficie de una esfera de radio $\sqrt{k}$. Para una exposición más amplia del tema, ver Myers y Montgomery [85a].



a) Contornos de $\sqrt{V[\hat{y}(\mathbf{x})]}$



b) La gráfica de la superficie de respuesta

Figura 11-21 Contornos de desviación estándar constante de la respuesta predicha para el DCC rotable, ejemplo 11-2.

Corridas centrales en el DCC

La elección de α en el DCC está dictada principalmente por la región de interés. Cuando esta región es una esfera, el diseño debe incluir corridas centrales para proporcionar una varianza razonablemente estable de la respuesta predicha. En general, se recomiendan de tres a cinco corridas centrales.

El diseño de Box-Behnken

Box y Behnken [13] han propuesto algunos diseños de tres niveles para ajustar superficies de respuesta. Estos diseños se forman combinando factoriales 2^k con diseños de bloques incompletos. Los diseños re-

Tabla 11-8 Diseño de Box-Behnken para tres variables

Corrida	x_1	x_2	x_3
1	−1	−1	0
2	−1	1	0
3	1	−1	0
4	1	1	0
5	−1	0	−1
6	−1	0	1
7	1	0	−1
8	1	0	1
9	0	−1	−1
10	0	−1	1
11	0	1	−1
12	0	1	1
13	0	0	0
14	0	0	0
15	0	0	0

sultantes suelen ser muy eficientes en términos del número requerido de corridas, y son rotables o casi rotables.

En la tabla 11-8 se muestra el diseño de Box-Behnken para tres variables. El diseño también se ilustra geométricamente en la figura 11-22. Observe que el diseño de Box-Behnken es un diseño esférico, con todos los puntos localizados en una esfera de radio $\sqrt{2}$. Asimismo, el diseño de Box-Behnken no contiene ningún punto en los vértices de la región cúbica creada por los límites superior e inferior de cada variable. Esto podría ser una ventaja cuando los puntos de los vértices del cubo representan combinaciones de los niveles de los factores cuya prueba es prohibitivamente costosa o imposible debido a restricciones físicas del proceso.

Región cuboidal de interés

Existen muchas situaciones en las que la región de interés es cuboidal en lugar de esférica. En estos casos, una variante útil del diseño central compuesto es el **diseño central compuesto con centros en las caras** o el **cubo con centros en las caras,** en el que $\alpha = 1$. En este diseño los puntos axiales o estrella se localizan en los centros de las caras del cubo, como se muestra en la figura 11-23 para $k = 3$. Esta variante del diseño central compuesto se usa en ocasiones debido a que sólo requiere tres niveles de cada factor, y en la prác-

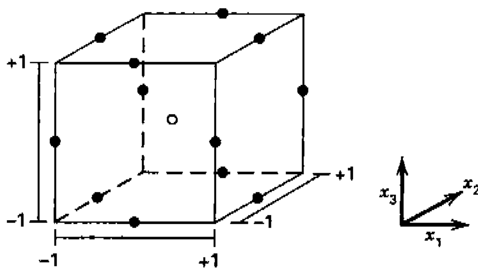


Figura 11-22 Diseño de Box-Behnken para tres factores.

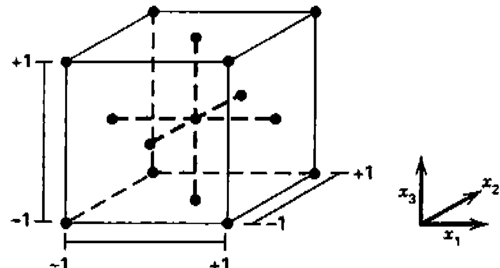
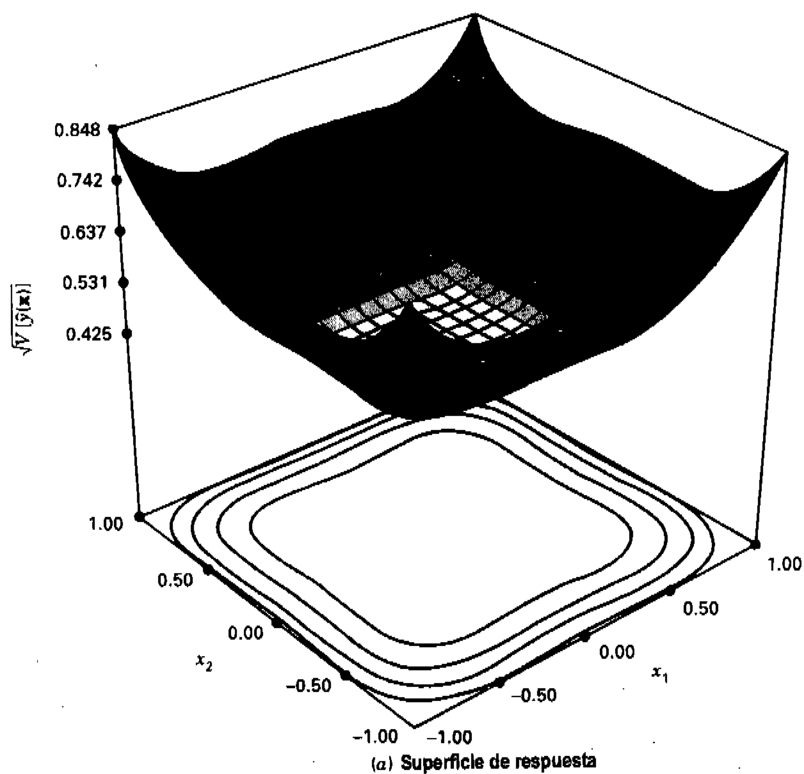
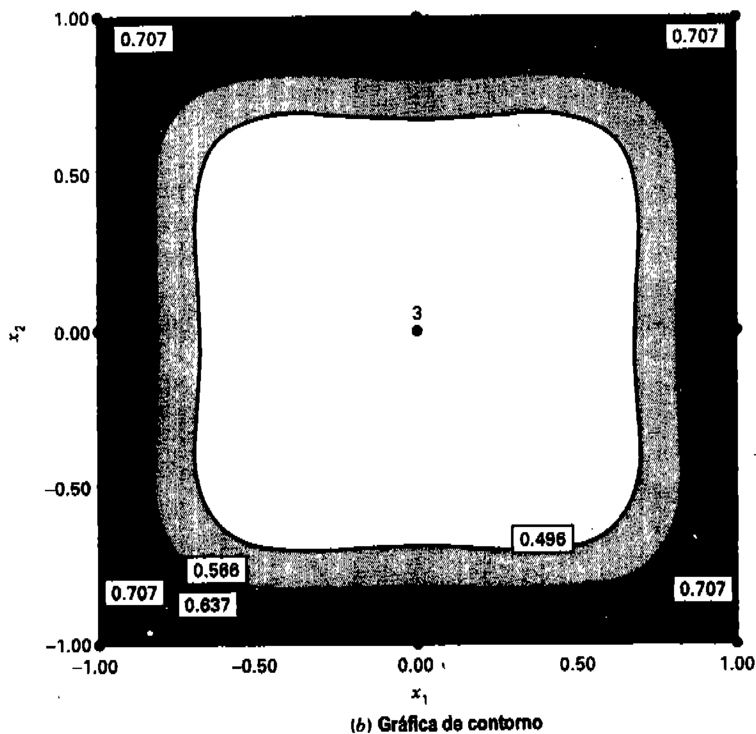


Figura 11-23 Diseño central compuesto con centros en las caras para $k = 3$.



(*a*) Superficie de respuesta



(*b*) Gráfica de contorno

Figura 11-24 Desviación estándar de la respuesta predicha $\sqrt{V[\hat{y}(\mathbf{x})]}$ para el cubo con centros en las caras con $k = 3$, $n_C = 3$ y $x_3 = 0$.

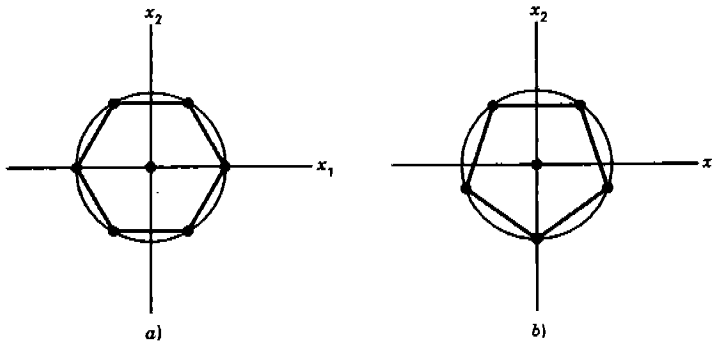


Figura 11-25 Diseños equirradiales para dos variables. *a*) Hexágono. *b*) Pentágono.

tica con frecuencia es difícil cambiar los niveles de los factores. Sin embargo, observe que los diseños centrales compuestos no son rotables.

El cubo con centros en las caras no requiere tantos puntos centrales como el DCC esférico. En la práctica, $n_C = 2$ o 3 es suficiente para proporcionar una buena varianza de predicción en toda la región experimental. Cabe señalar que en ocasiones se emplearán más corridas centrales para dar una estimación razonable del error experimental. En la figura 11-24 se muestra la raíz cuadrada de la varianza de predicción $\sqrt{V[\hat{y}(\mathbf{x})]}$ del cubo con centros en las caras para $k = 3$ con $n_C = 3$ puntos centrales ($x_3 = 0$). Observe que la desviación estándar de la respuesta predicha es razonablemente uniforme en una porción relativamente larga del espacio del diseño.

Otros diseños

Existen muchos otros diseños de superficie de respuesta que en ocasiones son útiles en la práctica. Para dos variables, podrían usarse diseños compuestos de puntos cuya separación en un círculo es igual y que forman polígonos regulares. Puesto que los puntos del diseño son equidistantes del origen, a estos arreglos con frecuencia se les llama **diseños equirradiales.**

Para $k = 2$, un diseño equirradial rotable se obtiene combinando $n_2 \geq 5$ puntos con una separación igual en un círculo con $n_1 \geq 1$ punto en el centro del círculo. Diseños de particular utilidad para $k = 2$ son el pentágono y el hexágono. Estos diseños se muestran en la figura 11-25. Otros diseños útiles incluyen el **diseño compuesto pequeño,** el cual consiste en un factorial fraccionado en el cubo de resolución III* (los efectos principales son alias de las interacciones de dos factores y ninguna de las interacciones de dos factores es alias entre sí) y las corridas axiales y centrales usuales, y la clase de los **diseños híbridos.** Estos diseños pueden ser de valor considerable cuando es importante reducir el número de corridas tanto como sea posible.

En la tabla 11-9 se muestra un diseño compuesto pequeño para $k = 3$ factores. Este diseño usa la fracción un medio estándar del diseño 2^3 en el cubo, ya que satisface los criterios de la resolución III*. El diseño tiene cuatro corridas en el cubo y seis corridas axiales, y debe incluir al menos un punto central. Por lo tanto, el diseño tiene un mínimo de $N = 11$ ensayos, y el modelo de segundo orden en $k = 3$ variables tiene $p = 10$ parámetros por estimar, por lo que se trata de un diseño muy eficiente con respecto al número de corridas. El diseño de la tabla 11-9 tiene $n_C = 4$ corridas centrales. Se seleccionó $\alpha = 1.73$ para obtener un diseño esférico debido a que el diseño compuesto pequeño no puede hacerse rotable.

En la tabla 11-10 se muestra un diseño híbrido para $k = 3$. Algunos de estos diseños tienen niveles irregulares, y esto puede ser un factor limitante para su aplicación. Sin embargo, se trata de diseños muy

Tabla 11-9 Diseño compuesto pequeño para $k = 3$ factores

Orden estándar	x_1	x_2	x_3
1	1.00	1.00	–1.00
2	1.00	–1.00	1.00
3	–1.00	1.00	1.00
4	–1.00	–1.00	–1.00
5	–1.73	0.00	0.00
6	1.73	0.00	0.00
7	0.00	–1.73	0.00
8	0.00	1.73	0.00
9	0.00	0.00	–1.73
10	0.00	0.00	1.73
11	0.00	0.00	0.00
12	0.00	0.00	0.00
13	0.00	0.00	0.00
14	0.00	0.00	0.00

pequeños, y poseen excelentes propiedades de la varianza de predicción. Para mayores detalles acerca de los diseños compuestos pequeños y los diseños híbridos, referirse a Myers y Montgomery [85a].

11-4.3 Formación de bloques en los diseños de superficie de respuesta

Cuando se usan diseños de superficie de respuesta, con frecuencia es necesario considerar la formación de bloques para eliminar las variables perturbadoras. Por ejemplo, este problema puede ocurrir cuando un diseño de segundo orden se ensambla secuencialmente a partir de un diseño de primer orden, como se ilustró en los ejemplos 11-1 y 11-2. Puede transcurrir tiempo considerable entre que se corre el modelo de primer orden y se corren los experimentos complementarios requeridos para construir un diseño de se-

Tabla 11-10 Diseño híbrido para $k = 3$ factores

Orden estándar	x_1	x_2	x_3
1	0.00	0.00	1.41
2	0.00	0.00	–1.41
3	–1.00	–1.00	0.71
4	1.00	–1.00	0.71
5	–1.00	1.00	0.71
6	1.00	1.00	0.71
7	1.41	0.00	–0.71
8	–1.41	0.00	–0.71
9	0.00	1.41	–0.71
10	0.00	–1.41	–0.71
11	0.00	0.00	0.00

gundo orden, y durante este tiempo las condiciones de prueba pueden cambiar, haciendo necesaria la formación de bloques.

Se dice que un diseño de superficie de respuesta se **forma de bloques ortogonales** si se divide en bloques tales que sus efectos no afecten las estimaciones de los parámetros del modelo de superficie de respuesta. Si se usa un diseño 2^k o 2^{k-p} como un diseño de superficie de respuesta de primer orden, pueden usarse los métodos del capítulo 7 para disponer las corridas en 2^r bloques. Los puntos centrales de estos diseños deberán asignarse por igual entre los bloques.

Para hacer la formación de bloques ortogonales de un diseño de segundo orden, deben satisfacerse dos condiciones. Si hay n_b observaciones en el bloque b-ésimo, entonces estas condiciones son

1. Cada bloque debe ser un diseño ortogonal de primer orden; es decir,

$$\sum_{u=1}^{n_b} x_{iu}x_{ju} = 0 \quad i \neq j = 0, 1, \ldots, k \qquad \text{para toda } b$$

donde x_{iu} y x_{ju} son los niveles de las variables i-ésima y j-ésima en la corrida u-ésima del experimento con $x_{0u} = 1$ para toda u.

2. La fracción de la suma de cuadrados total para cada variable con que contribuye cada bloque, debe ser igual a la fracción de las observaciones totales que están contenidas en el bloque; es decir,

$$\frac{\sum_{u=1}^{n_b} x_{iu}^2}{\sum_{u=1}^{N} x_{iu}^2} = \frac{n_b}{N} \qquad i = 1, 2, \ldots, k \qquad \text{para toda } b$$

donde N es el número de corridas del diseño.

Como un ejemplo de la aplicación de estas condiciones, considere un diseño central compuesto rotable en $k = 2$ variables con $N = 12$ corridas. Los niveles x_1 y x_2 de este diseño pueden escribirse en la matriz del diseño

$$\mathbf{D} = \begin{array}{c} \begin{matrix} x_1 & x_2 \end{matrix} \\ \left[\begin{array}{rr} -1 & -1 \\ 1 & -1 \\ -1 & 1 \\ 1 & 1 \\ 0 & 0 \\ 0 & 0 \\ 1.414 & 0 \\ -1.414 & 0 \\ 0 & 1.414 \\ 0 & -1.414 \\ 0 & 0 \\ 0 & 0 \end{array}\right] \end{array} \begin{array}{l} \left.\begin{array}{c} \\ \\ \\ \\ \\ \\ \end{array}\right\} \text{Bloque 1} \\ \left.\begin{array}{c} \\ \\ \\ \\ \\ \\ \end{array}\right\} \text{Bloque 2} \end{array}$$

Observe que el diseño se ha dispuesto en dos bloques, con el primer bloque consistiendo en la porción factorial del diseño más dos puntos centrales y el segundo bloque consistiendo en los puntos axiales más

dos puntos centrales adicionales. Es claro que la condición 1 se satisface; es decir, ambos bloques son diseños de primer orden ortogonales. Para investigar la condición dos, considere primero el bloque 1 y observe que

$$\sum_{u=1}^{n_1} x_{1u}^2 = \sum_{u=1}^{n_1} x_{2u}^2 = 4$$

$$\sum_{u=1}^{N} x_{1u}^2 = \sum_{u=1}^{N} x_{2u}^2 = 8 \qquad \text{y} \qquad n_1 = 6$$

Por lo tanto,

$$\frac{\sum_{u=1}^{n_1} x_{iu}^2}{\sum_{u=1}^{N} x_{iu}^2} = \frac{n_1}{N}$$

$$\frac{4}{8} = \frac{6}{12}$$

Así, la condición 2 se satisface en el bloque 1. Para el bloque 2 se tiene

$$\sum_{u=1}^{n_2} x_{1u}^2 = \sum_{u=1}^{n_2} x_{2u}^2 = 4 \qquad \text{y} \qquad n_2 = 6$$

Por lo tanto,

$$\frac{\sum_{u=1}^{n_2} x_{iu}^2}{\sum_{u=1}^{N} x_{iu}^2} = \frac{n_2}{N}$$

$$\frac{4}{8} = \frac{6}{12}$$

Puesto que la condición 2 también se satisface en el bloque 2, este diseño está formado de bloques ortogonales.

En general, el diseño central compuesto siempre puede construirse para hacer la formación de bloques ortogonales en dos bloques con el primer bloque consistiendo en n_F puntos factoriales más n_{CF} puntos centrales y el segundo bloque consistiendo en $n_A = 2k$ puntos axiales más n_{CA} puntos centrales. La primera condición de la formación de bloques ortogonales se cumplirá siempre independientemente del valor que se use para α en el diseño. Para que la segunda condición se cumpla,

$$\frac{\sum_{u}^{n_2} x_{iu}^2}{\sum_{u}^{n_1} x_{iu}^2} = \frac{n_A + n_{CA}}{n_F + n_{CF}} \tag{11-15}$$

El miembro izquierdo de la ecuación 11-15 es $2\alpha^2/n_F$, y después de sustituir esta cantidad, la ecuación para el valor de α que resultará en la formación de bloques ortogonales puede resolverse como

$$\alpha = \left[\frac{n_F(n_A + n_{CA})}{2(n_F + n_{CF})}\right]^{1/2} \tag{11-16}$$

Este valor de α no dará como resultado, en general, un diseño rotable o esférico. Si se requiere que el diseño también sea rotable, entonces $\alpha = (n_F)^{1/4}$ y

$$(n_F)^{1/2} = \frac{n_F(n_A + n_{CA})}{2(n_F + n_{CF})} \tag{11-17}$$

No siempre es posible encontrar un diseño que satisfaga exactamente la ecuación 11-17. Por ejemplo, si $k = 3$, $n_F = 8$ y $n_A = 6$, la ecuación 11-17 se reduce a

$$(8)^{1/2} = \frac{8(6+n_{CA})}{2(8+n_{CF})}$$

$$2.83 = \frac{48+8n_{CA}}{16+2n_{CF}}$$

Es imposible encontrar valores de n_{CA} y n_{CF} que satisfagan exactamente esta última ecuación. Sin embargo, observe que si $n_{CF} = 3$ y $n_{CA} = 2$, entonces el segundo miembro es

$$\frac{48+8(2)}{16+2(3)} = 2.91$$

por lo que el diseño se separa en bloques casi ortogonales. En la práctica podría relajarse un tanto el requerimiento de la rotabilidad o bien el de la formación de bloques ortogonales sin ninguna pérdida importante de información.

El diseño central compuesto es muy versátil en cuanto a su capacidad para incorporar la formación de bloques. Si k es lo suficientemente grande, la porción factorial del diseño puede dividirse en dos o más bloques. (El número de bloques factoriales debe ser una potencia de 2, con la porción axial formando un solo bloque.) En la tabla 11-11 se presentan varias disposiciones útiles de la formación de bloques para el diseño central compuesto.

Tabla 11-11 Algunos diseños centrales compuestos rotables y casi rotables que se separan en bloques ortogonales

k	2	3	4	5	5 $\frac{1}{2}$Rep.	6	6 $\frac{1}{2}$Rep.	7	7 $\frac{1}{2}$Rep.
Bloque(s) factorial(es)									
n_F	4	8	16	32	16	64	32	128	64
Número de bloques	1	2	2	4	1	8	2	16	8
Número de puntos en cada bloque	4	4	8	8	16	8	16	8	8
Número de puntos centrales en cada bloque	3	2	2	2	6	1	4	1	1
Número total de puntos en cada bloque	7	6	10	10	22	9	20	9	9
Bloque axial									
n_A	4	6	8	10	10	12	12	14	14
n_{CA}	3	2	2	4	1	6	2	11	4
Número total de puntos en el bloque axial	7	8	10	14	11	18	14	25	18
Número total de puntos N del diseño	14	20	30	54	33	90	54	169	80
Valores de α									
Separación en bloques ortogonales	1.4142	1.6330	2.0000	2.3664	2.0000	2.8284	2.3664	3.3636	2.8284
Rotabilidad	1.4142	1.6818	2.0000	2.3784	2.0000	2.8284	2.3784	3.3333	2.8284

Cabe destacar dos puntos importantes acerca del análisis de varianza cuando el diseño de superficie de respuesta se ha corrido en bloques. El primero se refiere al uso de los puntos centrales para calcular una estimación del error puro. Sólo los puntos centrales que se corren en el mismo bloque pueden considerarse como réplicas, por lo que el término del error puro sólo puede calcularse dentro de cada bloque. Si la variabilidad es consistente en todos los bloques, entonces estas estimaciones del error puro podrían agruparse. El segundo punto se refiere al efecto de bloque. Si el diseño se forma de bloques ortogonales en m bloques, la suma de cuadrados de los bloques es

$$SS_{\text{Bloques}} = \sum_{b=1}^{m} \frac{B_b^2}{n_b} - \frac{G^2}{N} \tag{11-18}$$

donde B_b es el total de las n_b observaciones en el bloque b-ésimo y G es el gran total de las N observaciones en los m bloques. Cuando los bloques no son exactamente ortogonales, puede usarse la prueba general de significación de la regresión (el método de la "suma de cuadrados extra") que se describió en el capítulo 10.

11-4.4 Diseños (óptimos) generados por computadora

Los diseños estándares de superficie de respuesta estudiados en las secciones anteriores, como el diseño central compuesto y el diseño de Box-Behnken y sus variantes (como el cubo con centros en las caras), son de uso generalizado porque son diseños bastante generales y flexibles. Si la región experimental es un cubo o una esfera, de manera típica existe un diseño de superficie de respuesta que será aplicable al problema. Sin embargo, ocasionalmente un experimentador se encuentra con una situación en la que el diseño estándar de superficie de respuesta puede no ser una elección obvia. Los **diseños generados por computadora** son una alternativa por considerar en estos casos.

Hay tres situaciones en las que puede ser apropiado algún tipo de diseño generado por computadora.

1. Una región experimental irregular. Si la región de interés del experimento no es un cubo o una esfera, los diseños estándares quizá no sean la mejor elección. Las regiones de interés irregulares ocurren con bastante frecuencia. Por ejemplo, un experimentador está investigando las propiedades de un adhesivo particular. El adhesivo se aplica a dos piezas y después se cura a una temperatura elevada. Los dos factores de interés son la cantidad de adhesivo aplicada y la temperatura de curado. En los rangos de estos dos factores, tomados como –1 a +1 en la escala de la variable codificada usual, el experimentador sabe que si se aplica muy poco adhesivo y la temperatura de curado es muy baja, las piezas no se pegarán satisfactoriamente. En términos de las variables codificadas, esto lleva a una **restricción** sobre las variables del diseño, por ejemplo

$$-1.5 \leq x_1 + x_2$$

donde x_1 representa la cantidad aplicada de adhesivo y x_2 la temperatura. Además, si la temperatura es demasiado elevada y se aplica mucho adhesivo, las piezas resultarán dañadas por la fatiga térmica o bien ocurrirá un pegado inadecuado. Por lo tanto, hay otra restricción sobre los niveles de los factores

$$x_1 + x_2 \leq 1$$

En la figura 11-26 se muestra la región experimental que resulta de aplicar estas restricciones. Observe que las restricciones eliminan de hecho dos de los vértices del cuadrado, produciendo una región experi-

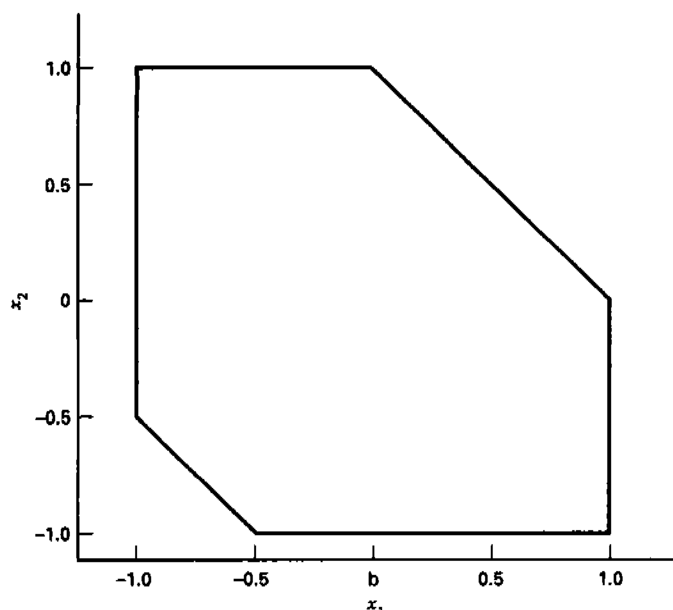


Figura 11-26 Región restringida del diseño en dos variables.

mental irregular (en ocasiones a estas regiones irregulares se les llama "latas abolladas"). No existe ningún diseño de superficie de respuesta estándar que se ajuste exactamente a esta región.

2. Un modelo no estándar. Por lo general, un experimentador elige un modelo de superficie de respuesta de primer o de segundo orden, consciente de que este **modelo empírico** es una aproximación del verdadero mecanismo subyacente. Sin embargo, en ocasiones el experimentador puede tener un conocimiento o idea especial acerca del proceso bajo estudio que puede sugerir un modelo no estándar. Por ejemplo, el modelo

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2$$
$$+ \beta_{112} x_1^2 x_2 + \beta_{1112} x_1^3 x_2 + \varepsilon$$

puede ser de interés. El experimentador estaría interesado en obtener un diseño eficiente para ajustar este modelo reducido de cuarto grado. Como otra ilustración, en ocasiones se encuentran problemas de superficie de respuesta en los que algunos de los factores del diseño son variables categóricas. No hay diseños de superficie de respuesta estándares para esta situación (referirse a Myers y Montgomery [85a] para un estudio de las variables categóricas en problemas de superficie de respuesta).

3. Requerimientos inusuales para el tamaño de la muestra. Ocasionalmente, un experimentador quizá necesite reducir el número de corridas requeridas en un diseño estándar de superficie de respuesta. Por ejemplo, suponga que se pretende ajustar un modelo de segundo orden en cuatro variables. El diseño central compuesto para esta situación requiere entre 28 y 30 corridas, dependiendo del número de puntos centrales seleccionados. Sin embargo, el modelo sólo tiene 15 términos. Si las corridas tienen un costo muy elevado o se llevan mucho tiempo, el experimentador querrá un diseño con menos ensayos. Aun cuando los diseños generados por computadora pueden usarse para este fin, por lo general se cuenta con enfoques mejores. Por ejemplo, puede construirse un diseño compuesto pequeño para cuatro factores

con 20 corridas, incluyendo cuatro puntos centrales, y también se cuenta con un diseño híbrido con apenas 16 corridas. Éstas son en general elecciones superiores al uso de un diseño generado por computadora para reducir el número de ensayos.

Gran parte del desarrollo de los diseños generados por computadora se deriva del trabajo de Kiefer [65a, b] y Kiefer y Wolfowitz [66] en la teoría de los **diseños optimales.** Por diseño optimal se entiende un diseño que es "mejor" con respecto a algún criterio. Se requieren programas de computadora para construir estos diseños. El enfoque usual es especificar un modelo, determinar la región de interés, seleccionar el número de corridas que deberán hacerse, especificar el criterio de optimalidad y después elegir los puntos del diseño de un conjunto de puntos *candidatos* que el experimentador consideraría usar. De manera típica, los puntos candidatos son una matriz de puntos distribuidos en la región factible del diseño.

Hay varios criterios de optimalidad populares. Quizá el de uso más generalizado es el criterio de optimalidad *D*. Se dice que un diseño es **optimal D** si

$$|(\mathbf{X}'\mathbf{X})^{-1}|$$

se minimiza. Ocurre que un diseño optimal *D* minimiza el volumen de la región de confianza conjunta para el vector de los coeficientes de regresión. Una medida de la eficiencia relativa del diseño 1 respecto del diseño 2 de acuerdo con el criterio *D* está dada por

$$D_e = \left(\frac{|(\mathbf{X}_2'\mathbf{X}_2)^{-1}|}{|(\mathbf{X}_1'\mathbf{X}_1)^{-1}|} \right)^{1/p} \tag{11-19}$$

donde $\mathbf{X}_1$ y $\mathbf{X}_2$ son las matrices **X** de los dos diseños y *p* es el número de parámetros del modelo.

El criterio de **optimalidad** *A* sólo se ocupa de las varianzas de los coeficientes de regresión. Un diseño es optimal *A* si minimiza la suma de los elementos de la diagonal principal de $(\mathbf{X}'\mathbf{X})^{-1}$ [a ésta se le llama la **traza** de $(\mathbf{X}'\mathbf{X})^{-1}$, denotada generalmente como $\text{tr}(\mathbf{X}'\mathbf{X})^{-1}$]. Por lo tanto, un diseño optimal *A* minimiza la suma de las varianzas de los coeficientes de regresión.

Puesto que muchos experimentos de superficie de respuesta se refieren a la predicción de la respuesta, los **criterios de la varianza de predicción** son de gran interés práctico. Quizás el más popular de estos criterios sea el **criterio de optimalidad** ***G***. Se dice que un diseño es optimal *G* si minimiza la varianza de predicción escalada máxima en la región del diseño. Es decir, si el valor máximo de

$$\frac{NV[\hat{y}(\mathbf{x})]}{\sigma^2}$$

en la región del diseño es un mínimo, donde *N* es el número de puntos del diseño. Si el modelo tiene *p* parámetros, la eficiencia *G* de un diseño es precisamente

$$G_e = \frac{p}{\text{máx}\,\dfrac{NV[\hat{y}(\mathbf{x})]}{\sigma^2}} \tag{11-20}$$

El **criterio** ***V*** considera la varianza de predicción en un *conjunto* de puntos de interés en la región del diseño, por ejemplo $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m$. El conjunto de puntos podría ser el conjunto de candidatos del que se seleccionó el diseño, o podría ser alguna otra colección de puntos que tienen un significado específico para el experimentador. Un diseño que minimiza la varianza de predicción *promedio* en este conjunto de *m* puntos es un **diseño optimal** ***V***.

En conjunto, a los criterios de diseño que se han venido estudiando suele llamárseles **criterios de optimalidad alfabética.** Existen algunas situaciones en las que el diseño optimal alfabético se conoce o bien

puede construirse analíticamente. Un buen ejemplo es el diseño 2^k, que es optimal *D, A, G* y *V* para ajustar el modelo de primer orden en *k* variables o para ajustar el modelo de primer orden con interacción. Sin embargo, en la mayoría de los casos el diseño optimal no se conoce y debe emplearse un algoritmo basado en computadora para encontrar un diseño. Muchos paquetes de software de estadística que soportan experimentos diseñados cuentan con esta capacidad. La mayoría de los procedimientos para construir diseños se basan en el **algoritmo de intercambio.** En esencia, el experimentador selecciona una matriz de puntos candidatos y un diseño inicial (quizá al azar) a partir de este conjunto de puntos. Entonces el algoritmo intercambia los puntos que están en la matriz, pero no en el diseño, con los puntos que están actualmente en el diseño, en un esfuerzo por mejorar el criterio de optimalidad seleccionado. Debido a que no se evalúan explícitamente todos los diseños posibles, no hay garantía de que se ha encontrado un diseño optimal, pero el procedimiento de intercambio suele asegurar que se obtiene un diseño que está "cerca" del optimal. Algunas implementaciones repiten varias veces el proceso de construcción del diseño, empezando con diseños iniciales diferentes, para incrementar la posibilidad de que se obtendrá un diseño final que esté muy cerca del optimal.

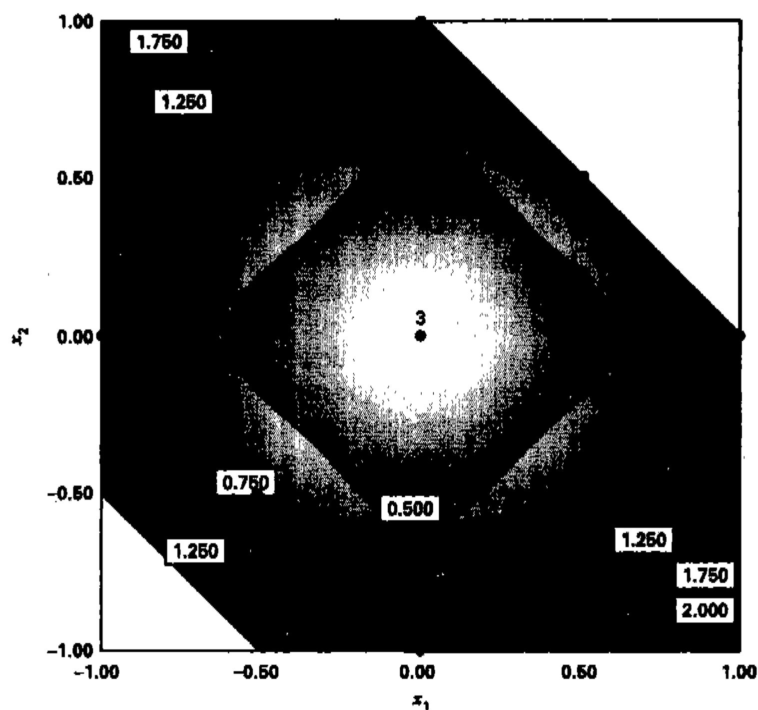
Para ilustrar algunas de estas ideas, considere el experimento del adhesivo expuesto anteriormente y que llevó a la región experimental irregular de la figura 11-26. Suponga que la respuesta de interés es la fuerza de desprendimiento y que quiere ajustarse un modelo de segundo orden para esta respuesta. En la figura 11-27*a* se muestra un diseño central compuesto con cuatro puntos centrales (12 corridas en total) inscrito dentro de esta región. Se trata de un diseño que no es rotable, pero es el DCC más grande que puede ajustarse dentro del espacio del diseño. Para este diseño $|(\mathbf{X}'\mathbf{X})^{-1}| = 1.852$ E–2, y la traza de $(\mathbf{X}'\mathbf{X})^{-1}$ es 6.375. En la figura 11-27*a* también se muestran los contornos de desviación estándar constante de la respuesta predicha, calculada suponiendo que $\sigma = 1$. En la figura 11-27*b* se muestra la gráfica de superficie de respuesta correspondiente.

En la figura 11-28*a* y en la tabla 11-12 se muestra un diseño optimal *D* de 12 corridas para este problema, generado con el paquete de software *Design-Expert*. Para este diseño, $|(\mathbf{X}'\mathbf{X})^{-1}| = 2.153$ E–4. Observe que el criterio *D* es considerablemente mejor para este diseño que el DCC inscrito. La eficiencia relativa del DCC inscrito con respecto al diseño optimal *D* es

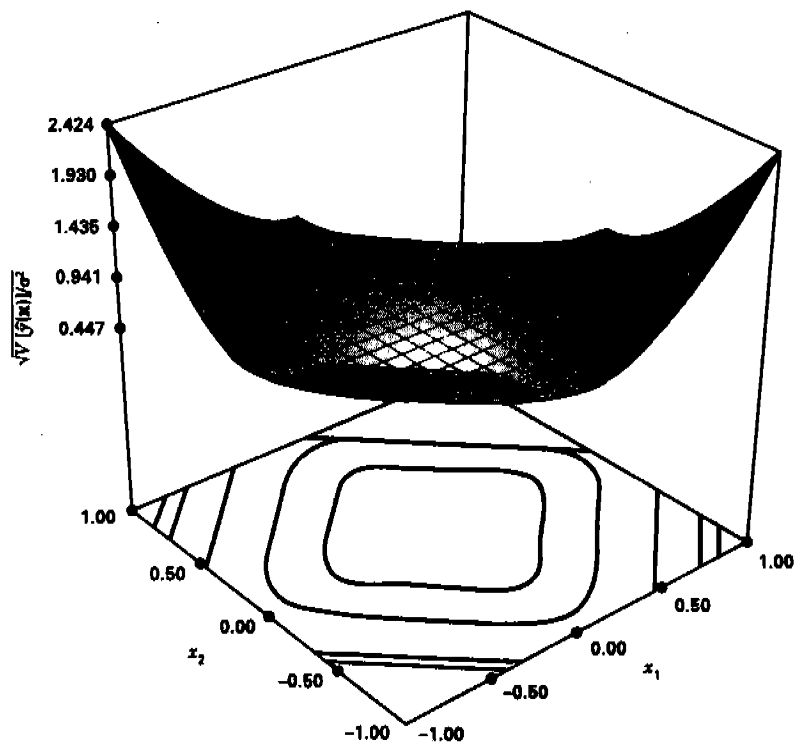
$$D_e = \left(\frac{|(\mathbf{X}_2'\mathbf{X}_2)^{-1}|}{|(\mathbf{X}_1'\mathbf{X}_1)^{-1}|}\right)^{1/p} = \left(\frac{0.0002153}{0.01852}\right)^{1/6} = 0.476$$

Es decir, el DCC inscrito tiene una eficiencia de sólo 47.6% que la del diseño optimal *D*. Esto implica que tendrían que hacerse $1/0.476 = 2.1$ réplicas del DCC (o aproximadamente el doble) para tener la misma precisión de la estimación de los coeficientes de regresión que la que se consigue con el diseño optimal *D*. La traza de $(\mathbf{X}'\mathbf{X})^{-1}$ es 2.516 para el diseño optimal *D*, lo cual indica que la suma de las varianzas de los coeficientes de regresión es considerablemente más pequeña para este diseño que para el DCC. En las figuras 11-28*a* y *b* se muestran también los contornos de desviación estándar constante de la respuesta predicha y la gráfica de la superficie de respuesta asociada (suponiendo que $\sigma = 1$). En general, los contornos de la desviación estándar de la predicción son más bajos para el diseño optimal *D* que para el DCC inscrito, particularmente cerca de los límites de la región de interés, donde el DCC inscrito no incluye ninguno de los puntos del diseño.

En la figura 11-29*a* se muestra un tercer diseño, creado al tomar las dos réplicas de los vértices de la región en el diseño optimal *D* y pasarlas al centro del diseño. Esto podría ser una idea útil, ya que la figura 11-28*b* muestra que la desviación estándar de la respuesta predicha se incrementa ligeramente cerca del centro de la región del diseño para el diseño optimal *D*. En la figura 11-29*a* se muestran también los contornos de desviación estándar constante de la predicción para este diseño optimal *D modificado*, y en la fi-

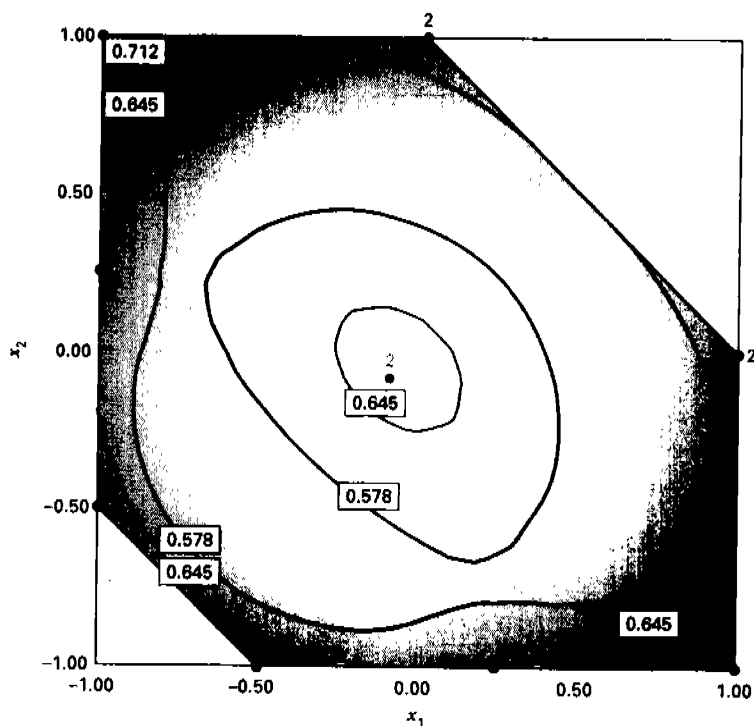


(*a*) El diseño y los contornos de $\sqrt{V[\hat{y}(\mathbf{x})]/\sigma^2}$ constante

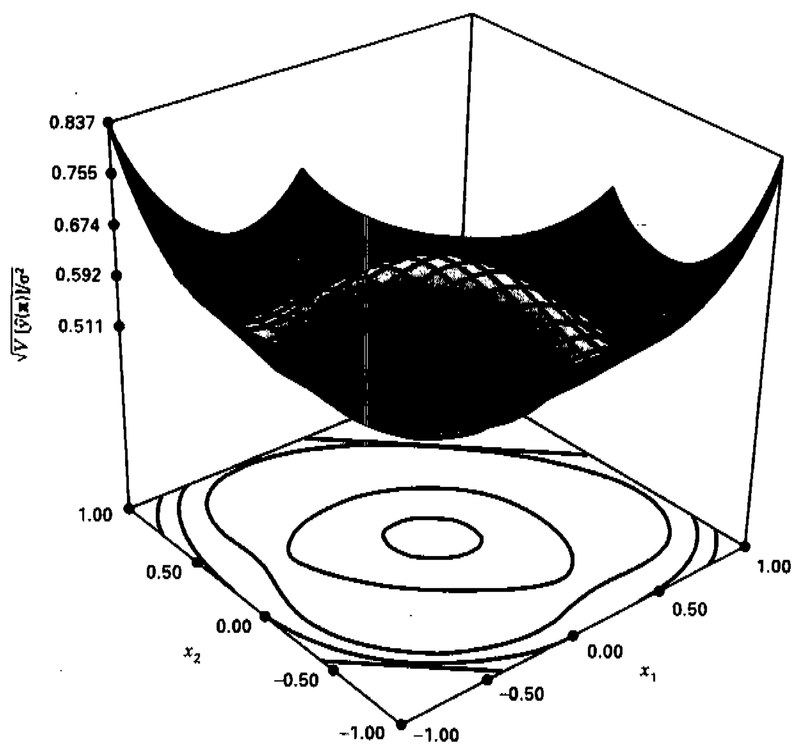


(*b*) La gráfica de superficie de respuesta

Figura 11-27 Un diseño central compuesto inscrito para la región restringida del diseño de la figura 11-26.



(*a*) El diseño y los contornos de $\sqrt{V[\hat{y}(\mathbf{x})]/\sigma^2}$ constante



(*b*) La gráfica de superficie de respuesta

Figura 11-28 Un diseño optimal *D* para la región restringida del diseño de la figura 11-26.

Tabla 11-12 Diseño optimal D para la región restringida de la figura 11-26

Orden estándar	x_1	x_2
1	−0.50	−1.00
2	1.00	0.00
3	−0.08	−0.08
4	−1.00	1.00
5	1.00	−1.00
6	0.00	1.00
7	−1.00	0.25
8	0.25	−1.00
9	−1.00	−0.50
10	1.00	0.00
11	0.00	1.00
12	−0.08	−0.08

gura 11-29*b* se muestra la gráfica de la superficie de respuesta. El criterio D para este diseño es $|(\mathbf{X}'\mathbf{X})^{-1}|$ = 3.71 E–4, y la eficiencia relativa es

$$D_e = \left(\frac{|(\mathbf{X}_2'\mathbf{X}_2)^{-1}|}{|(\mathbf{X}_1'\mathbf{X}_1)^{-1}|}\right)^{1/p} = \left(\frac{0.0002153}{0.000371}\right)^{1/6} = 0.91$$

Es decir, este diseño es casi tan eficiente como el diseño optimal D. La traza de $(\mathbf{X}'\mathbf{X})^{-1}$ es 2.448 para este diseño, un valor ligeramente mayor que el que se obtuvo para el diseño optimal D. Los contornos de desviación estándar constante de la predicción para este diseño dan la impresión visual de ser al menos tan buenos como los del diseño optimal D, particularmente en el centro de la región.

Los diseños generados por computadora con base en los criterios de optimalidad alfabética pueden ser ciertamente útiles en situaciones en las que la región experimental no es ni esférica ni cuboidal. Sin embargo, no son sustitutos de los diseños estándares en la mayoría de los problemas. Los diseños optimales alfabéticos se generan apegándose estrictamente a un solo criterio y, como se señaló al principio de la sección 11-4, donde se enlistaron varios criterios para diferentes diseños, incluyen varios que son de carácter un tanto cualitativo o subjetivo. En problemas experimentales reales, por lo general hay muchos criterios que es necesario evaluar para seleccionar un diseño. Para un estudio más amplio de este tema, referirse a Myers y Montgomery [85a, capítulo 8].

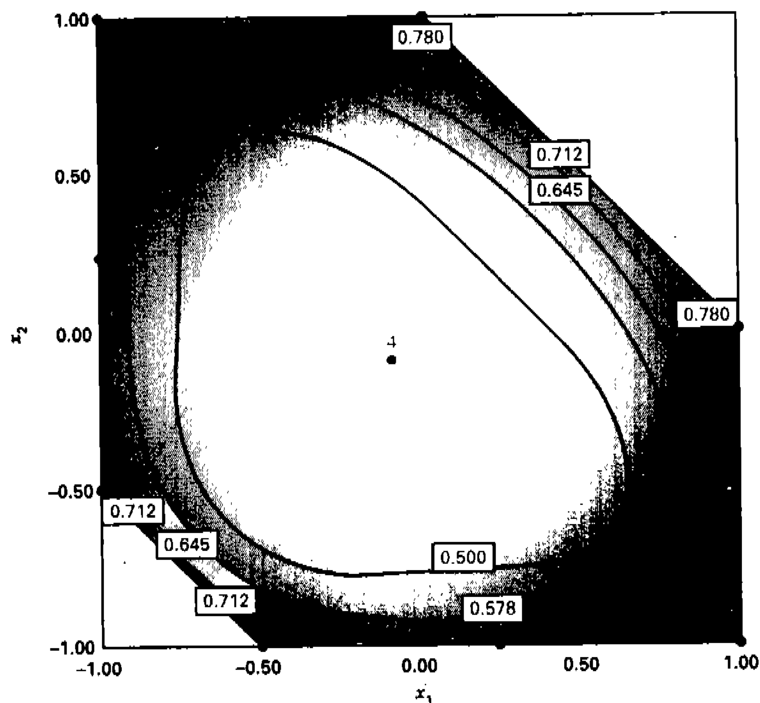
11-5 EXPERIMENTOS CON MEZCLAS

En las secciones anteriores se presentaron diseños de superficie de respuesta para aquellas situaciones en las que los niveles de cada factor son independientes de los niveles de otros factores. En los **experimentos con mezclas,** los factores son los componentes o ingredientes de una mezcla y, por consiguiente, sus niveles no son independientes. Por ejemplo, si $x_1, x_2, ..., x_p$ denota las proporciones de p componentes de una mezcla, entonces

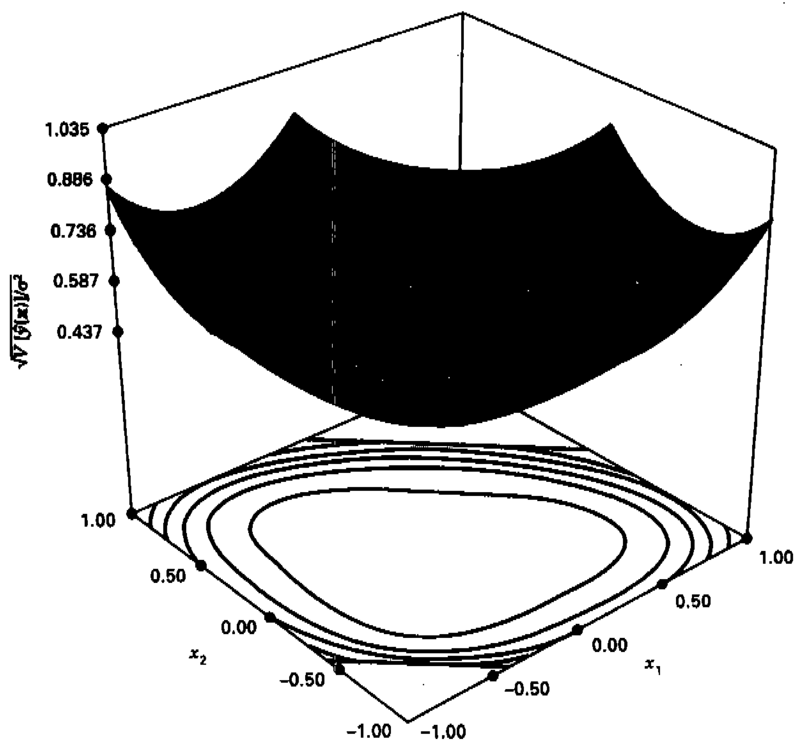
$$0 \le x_i \le 1 \qquad i = 1, 2, \ldots, p$$

y

$$x_1 + x_2 + \cdots + x_p = 1 \qquad \text{(es decir, 100\%)}$$

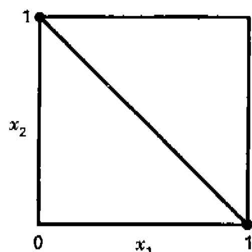


(a) El diseño y los contornos de $\sqrt{V[\hat{y}(\mathbf{x})]/\sigma^2}$ constante



(b) La gráfica de superficie de respuesta

Figura 11-29 Un diseño optimal D modificado para la región restringida del diseño de la figura 11-26.



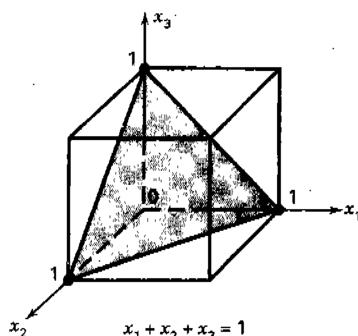


Figura 11-30 Espacio de los factores restringidos para mezclas con *a*) $p = 2$ componentes y *b*) $p = 3$ componentes.

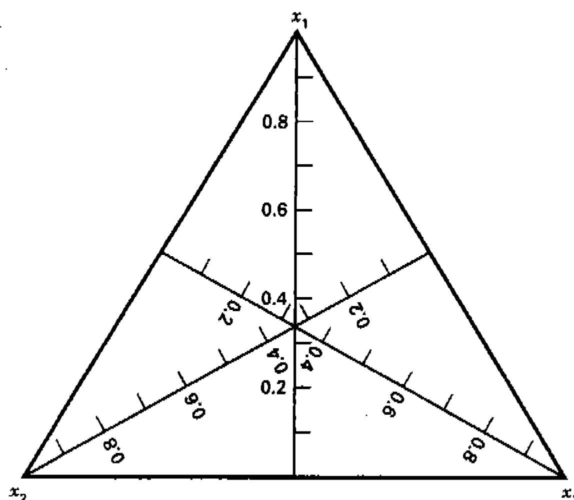


Figura 11-31 Sistema coordenado trilineal.

Estas restricciones se ilustran gráficamente en la figura 11-30 para $p = 2$ y $p = 3$ componentes. Para dos componentes, el espacio de los factores del diseño incluye todos los valores de los dos componentes que están sobre el segmento de recta $x_1 + x_2 = 1$, con cada componente siendo acotado por 0 y 1. Con tres componentes, el espacio de la mezcla es un triángulo con vértices que corresponden a las formulaciones que son **mezclas puras** (mezclas que son 100% de un solo componente).

Cuando hay tres componentes en la mezcla, la región experimental restringida puede representarse convenientemente en **papel milimétrico trilineal,** como se muestra en la figura 11-31. Cada uno de los tres lados de la gráfica de la figura 11-31 representa una mezcla que no contiene nada de alguno de los tres componentes (el componente indicado en el vértice opuesto). Las nueve líneas de graduación en cada dirección marcan incrementos de 10% en el componente respectivo.

Los **diseños símplex** se usan para estudiar los efectos de los componentes de una mezcla sobre la variable de respuesta. Un **diseño símplex reticular** $\{p, m\}$ para p componentes consta de los puntos definidos por los siguientes arreglos de las coordenadas: las proporciones asumidas por cada componente toman los $m + 1$ valores que están separados por una distancia igual de 0 a 1,

$$x_i = 0, \frac{1}{m}, \frac{2}{m}, \cdots, 1 \qquad i = 1, 2, \ldots, p \tag{11-21}$$

y se usan todas las combinaciones posibles (mezclas) de las proporciones de la ecuación 11-21. Como un ejemplo, sean $p = 3$ y $m = 2$. Entonces

$$x_i = 0, \tfrac{1}{2}, 1 \qquad i = 1, 2, 3$$

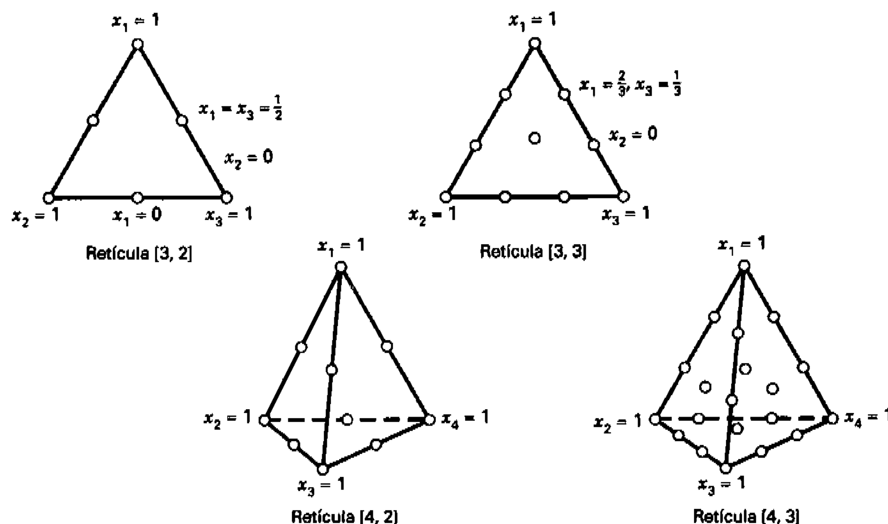


Figura 11-32 Algunos diseños símplex reticulares para $p = 3$ y $p = 4$ componentes.

y el diseño símplex reticular consta de las seis corridas siguientes:

$$(x_1, x_2, x_3) = (1, 0, 0), (0, 1, 0), (0, 0, 1), (\tfrac{1}{2}, \tfrac{1}{2}, 0), (\tfrac{1}{2}, 0, \tfrac{1}{2}), (0, \tfrac{1}{2}, \tfrac{1}{2})$$

En la figura 11-32 se ilustra este diseño. Los tres vértices (1, 0, 0), (0, 1, 0) y (0, 0, 1) son las mezclas puras, mientras que los puntos $(\frac{1}{2}, \frac{1}{2}, 0)$, $(\frac{1}{2}, 0, \frac{1}{2})$ y $(0, \frac{1}{2}, \frac{1}{2})$ son mezclas binarias o mezclas de dos componentes localizadas en los puntos medios de los tres lados del triángulo. En la figura 11-32 se muestran también los diseños símplex reticulares {3, 3}, {4, 2} y {4, 3}. En general, el número de puntos en un diseño símplex reticular $\{p, m\}$ es

$$N = \frac{(p+m-1)!}{m!(p-1)!}$$

Una alternativa del diseño símplex reticular es el **diseño símplex de centroide.** En un diseño símplex de centroide con p componentes, hay $2^p - 1$ puntos, que corresponden a las p permutaciones de (1, 0, 0, ..., 0), las $\binom{p}{2}$ permutaciones de $(\frac{1}{2}, \frac{1}{2}, 0, ..., 0)$, las $\binom{p}{3}$ permutaciones de $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0, ..., 0)$, ..., y el centroide global $(\frac{1}{p}, \frac{1}{p}, ..., \frac{1}{p})$. En la figura 11-33 se muestran algunos diseños símplex de centroide.

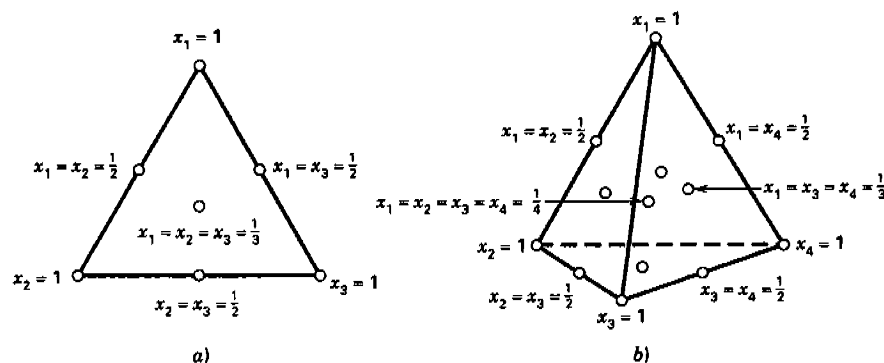


Figura 11-33 Diseños símplex de centroide con $a)$ $p = 3$ componentes y $b)$ $p = 4$ componentes.

Una crítica a los diseños símplex descritos antes es que la mayoría de las corridas ocurren en la frontera de la región y, por consiguiente, incluyen sólo $p-1$ de los p componentes. Suele ser deseable aumentar el diseño símplex reticular o de centroide con puntos adicionales en el interior de la región donde las mezclas estarán formadas por la totalidad de los p componentes. Para un estudio más amplio, ver Cornell [33] y Myers y Montgomery [85a].

Los modelos para mezclas difieren de los polinomios usuales empleados en los diseños de superficie de respuesta debido a la restricción $\Sigma x_i = 1$. Las formas estándares de los modelos para mezclas que se usan ampliamente son

Lineal:

$$E(y)=\sum_{i=1}^{p} \beta_i x_i \tag{11-22}$$

Cuadrático:

$$E(y)=\sum_{i=1}^{p} \beta_i x_i + \sum\sum_{i<j}^{p} \beta_{ij} x_i x_j \tag{11-23}$$

Cúbico completo:

$$\begin{aligned} E(y)=&\sum_{i=1}^{p} \beta_i x_i + \sum\sum_{i<j}^{p} \beta_{ij} x_i x_j \\ &+\sum\sum_{i<j}^{p} \delta_{ij} x_i x_j (x_i - x_j) \\ &+\sum\sum\sum_{i<j<k} \beta_{ijk} x_i x_j x_k \end{aligned} \tag{11-24}$$

Cúbico especial:

$$\begin{aligned} E(y)=&\sum_{i=1}^{p} \beta_i x_i + \sum\sum_{i<j}^{p} \beta_{ij} x_i x_j \\ &+\sum\sum\sum_{i<j<k} \beta_{ijk} x_i x_j x_k \end{aligned} \tag{11-25}$$

Los términos de estos modelos tienen interpretaciones relativamente simples. En las ecuaciones 11-22 a 11-25, el parámetro β_i representa la respuesta esperada para la mezcla pura $x_i = 1$ y $x_j = 0$ cuando $j \neq i$. A la porción $\Sigma_{i=1}^{p} \beta_i x_i$ se le llama **porción de mezcla lineal.** Cuando hay curvatura derivada de una mezcla no lineal entre pares de componentes, los parámetros β_{ij} representan una **mezcla sinérgica** o bien **antagónica.** Los términos de órdenes superiores suelen ser necesarios en los modelos para mezclas porque 1) los fenómenos estudiados pueden ser complejos y 2) la región experimental con frecuencia es la región de operabilidad completa y, en consecuencia, es grande y requiere un modelo elaborado.

EJEMPLO 11-3 ····························

Una mezcla de tres componentes

Cornell [33] describe el experimento con una mezcla en el que se combinaron tres componentes —polietileno (x_1), poliestireno (x_2) y polipropileno (x_3)— para hilar una fibra que se usará en cortinas. La variable de respuesta de interés es la elongación del hilo en kilogramos de fuerza aplicada. Se usa un diseño símplex

Tabla 11-13 El diseño símplex reticular {3, 2} para el problema de la elongación del hilo

Punto del diseño	Proporciones de los componentes x_1	x_2	x_3	Valores observados de la elongación	Valor promedio de la elongación ($\bar{y}$)
1	1	0	0	11.0, 12.4	11.7
2	$\frac{1}{2}$	$\frac{1}{2}$	0	15.0, 14.8, 16.1	15.3
3	0	1	0	8.8, 10.0	9.4
4	0	$\frac{1}{2}$	$\frac{1}{2}$	10.0, 9.7, 11.8	10.5
5	0	0	1	16.8, 16.0	16.4
6	$\frac{1}{2}$	0	$\frac{1}{2}$	17.7, 16.4, 16.6	16.9

reticular para estudiar el producto. El diseño y las respuestas observadas se muestran en la tabla 11-13. Observe que todos los puntos del diseño incluyen **mezclas puras** o **binarias**; es decir, únicamente se usan a lo sumo dos de los tres componentes en cualquier formulación del producto. También se corren réplicas de las observaciones, con dos réplicas de cada una de las mezclas puras y tres réplicas de cada una de las mezclas binarias. La desviación estándar del error puede estimarse a partir de estas réplicas de las observaciones como $\hat{\sigma} = 0.85$. Cornell ajusta el polinomio de segundo grado de la mezcla a los datos, obteniendo

$$\hat{y} = 11.7x_1 + 9.4x_2 + 16.4x_3 + 19.0x_1x_2 + 11.4x_1x_3 - 9.6x_2x_3$$

Puede demostrarse que este modelo es una representación adecuada de la respuesta. Observe que como $\hat{\beta}_3 > \hat{\beta}_1 > \hat{\beta}_2$, se concluiría que el componente 3 (polipropileno) produce el hilo con la elongación máxima. Además, puesto que $\hat{\beta}_{12}$ y $\hat{\beta}_{13}$ son positivos, la mezcla de los componentes 1 y 2 o de los componentes 1 y 3 produce valores más altos de la elongación de los que se esperarían si nos limitáramos a promediar las elongaciones de las mezclas puras. Se trata de un ejemplo de los efectos de mezclado "sinérgicos". Los componentes 2 y 3 tienen efectos de mezclado antagónicos, ya que $\hat{\beta}_{23}$ es negativa.

En la figura 11-34 se grafican los contornos de la elongación, lo cual puede ser de utilidad para interpretar los resultados. Al examinar la figura, se observa que si se desea la elongación máxima, deberá elegirse la mezcla de los componentes 1 y 3, la cual está formada por aproximadamente 80% del componente 3 y 20% del componente 1.

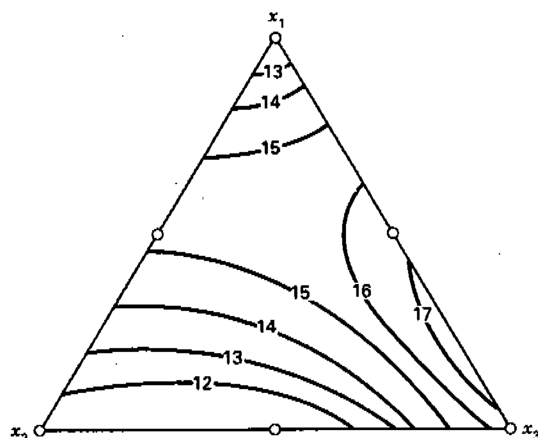


Figura 11-34 Contornos de la elongación estimada del hilo constante en el modelo de segundo orden para la mezcla del ejemplo 11-3.

Se señaló ya que los diseños símplex reticular y símplex de centroide son **diseños de puntos frontera.** Si el experimentador quiere hacer predicciones acerca de las propiedades de mezclas completas, sería muy deseable contar con más corridas en el interior del símplex. Se recomienda aumentar los diseños símplex ordinarios con **corridas axiales** y el centroide global (si el centroide no es ya un punto del diseño).

El **eje del componente *i*** es la recta o rayo que se extiende del punto base $x_i = 0, x_j = 1/(p-1)$, para toda $j \neq i$, al vértice opuesto donde $x_i = 1, x_j = 0$ para toda $j \neq i$. El punto base siempre se localizará en el centroide de la frontera de $(p-2)$ dimensiones del diseño símplex que está opuesto al vértice $x_i = 1, x_j = 0$ para toda $j \neq i$. [A la frontera se le llama en ocasiones el $(p-2)$-llano.] La longitud del eje del componente es una unidad. Los **puntos axiales** se sitúan sobre los ejes de los componentes a una distancia Δ del centroide. El valor máximo de Δ es $(p-1)/p$. Se recomienda que las corridas axiales se coloquen a la mitad entre el centroide del diseño símplex y cada vértice para que $\Delta = (p-1)/2p$. En ocasiones a estos puntos se les llama **mezclas de verificación axial,** porque es una práctica común excluirlas cuando se ajusta el modelo preliminar de la mezcla y usar después las respuestas en estos puntos axiales para verificar la adecuación del ajuste del modelo preliminar.

En la figura 11-35 se muestra el diseño símplex reticular {3, 2} aumentado con los puntos axiales. Este diseño tiene 10 puntos, con cuatro de ellos en el interior del diseño símplex. La retícula símplex {3, 3} soportará el ajuste del modelo cúbico completo, mientras que la retícula símplex aumentada no lo hará; sin embargo, la retícula símplex aumentada permitirá al experimentador ajustar el modelo cúbico especial o agregar al modelo cuadrático términos especiales de cuarto orden, como $\beta_{1233}x_1x_2x_3^2$. La retícula símplex aumentada es superior para estudiar la respuesta de mezclas completas en el sentido de que puede detectar y modelar la curvatura en el interior del triángulo que no puede tomarse en consideración por los términos del modelo cúbico completo. La retícula símplex aumentada tiene más potencia para detectar la falta de ajuste que la retícula {3, 3}. Esto es de particular utilidad cuando el experimentador no está seguro acerca del modelo apropiado que debe usar y también planea construir un modelo secuencialmente empezando con un polinomio simple (quizá de primer orden), probar el modelo para la

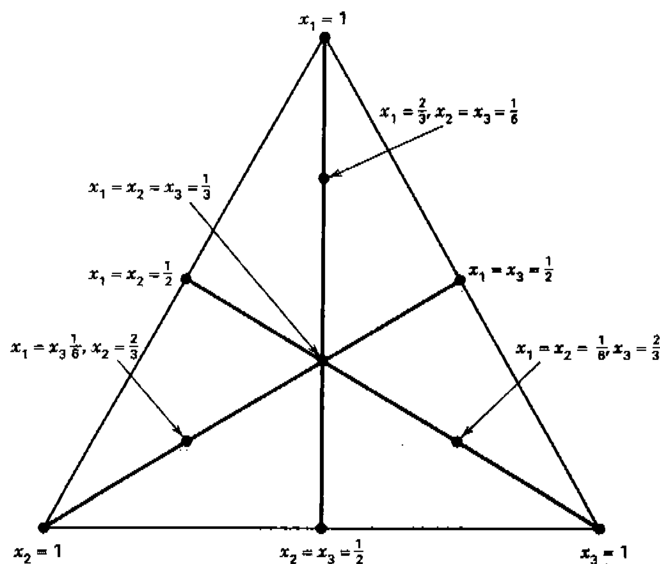


Figura 11-35 Un diseño símplex reticular aumentado.

falta de ajuste, después aumentar el modelo con términos de órdenes superiores, probar el nuevo modelo para la falta de ajuste y así sucesivamente.

En algunos problemas de mezclas surgen **restricciones** sobre los componentes individuales. Las restricciones sobre la frontera inferior de la forma

$$l_i \leq x_i \leq 1 \qquad i = 1, 2, \ldots, p$$

son muy comunes. Cuando sólo están presentes restricciones sobre la frontera inferior, la región factible del diseño sigue siendo un diseño símplex, pero se inscribe dentro de la región del símplex original. Esta situación puede simplificarse mediante la introducción de **pseudocomponentes,** definidos como

$$x_i' = \frac{x_i - l_i}{\left(1 - \sum_{j=1}^{p} l_j\right)} \qquad (11\text{-}26)$$

con $\Sigma_{j=1}^{p} l_j < 1$. Entonces

$$x_1' + x_2' + \cdots + x_p' = 1$$

por lo que el uso de pseudocomponentes permite utilizar diseños tipo símplex cuando las fronteras inferiores forman parte de la situación experimental. Las formulaciones especificadas por el diseño símplex para los pseudocomponentes se transforman en formulaciones para los componentes originales invirtiendo la transformación de la ecuación 11-26. Es decir, si x_i' es el valor asignado al pseudocomponente i-ésimo en una de las corridas del experimento, el componente i-ésimo de la mezcla original es

$$x_i = l_i + \left(1 - \sum_{j=1}^{p} l_j\right) x_i' \qquad (11\text{-}27)$$

Cuando los componentes tienen restricciones tanto sobre la frontera superior como la inferior, la región factible deja de ser un diseño símplex; será, en cambio, un politopo irregular. Puesto que la región experimental no tiene una forma "estándar", los **diseños generados por computadora** son muy útiles para este tipo de problemas de mezclas.

EJEMPLO 11-4 ··

Formulación de una pintura

Un experimentador está intentando optimizar la formulación de una pintura automotriz de recubrimiento total. Se trata de productos complejos que tienen requerimientos de desempeño muy específicos. El cliente quiere, en particular, que la dureza Knoop exceda de 25 y que el porcentaje de sólidos esté abajo de 30. El recubrimiento total es una mezcla de tres componentes, que consiste en un monómero (x_1), un entrelazador (x_2) y una resina (x_3). Existen restricciones sobre las proporciones de los componentes:

$$x_1 + x_2 + x_3 = 100$$
$$5 \leq x_1 \leq 25$$
$$25 \leq x_2 \leq 40$$
$$50 \leq x_3 \leq 70$$

El resultado es la región de experimentación restringida ilustrada en la figura 11-36. Puesto que la región de interés no es símplex, se usará un diseño optimal D para este problema. Suponiendo que posiblemente ambas respuestas serán modeladas con un modelo cuadrático de una mezcla, el diseño optimal D ilustra-

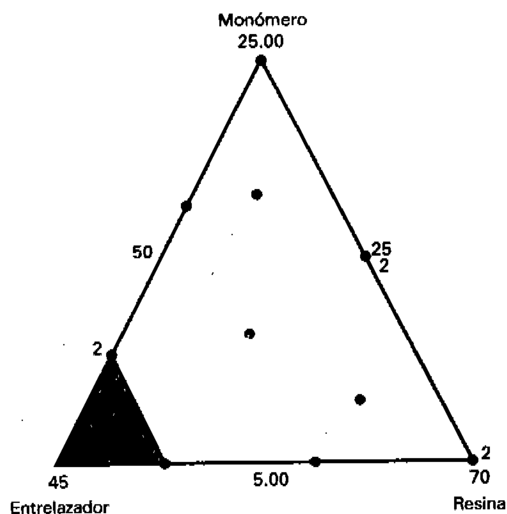


Figura 11-36 La región experimental restringida para el problema de la formulación de la pintura del ejemplo 11-4 (mostrada en la escala real del componente).

Tabla 11-14 Diseño optimal *D* para el problema de la formulación de la pintura del ejemplo 11-4

Orden estándar	Corrida	Monómero x_1	Entrelazador x_2	Resina x_3	Dureza y_1	Sólidos y_2
1	2	17.50	32.50	50.00	29	9.539
2	1	10.00	40.00	50.00	26	27.33
3	4	15.00	25.00	60.00	17	29.21
4	13	25.00	25.00	50.00	28	30.46
5	7	5.00	25.00	70.00	35	74.98
6	3	5.00	32.50	62.50	31	31.5
7	6	11.25	32.50	56.25	21	15.59
8	11	5.00	40.00	55.00	20	19.2
9	10	18.13	28.75	53.33	29	23.44
10	14	8.13	28.75	63.13	25	32.49
11	12	25.00	25.00	50.00	19	23.01
12	9	15.00	25.00	60.00	14	41.46
13	5	10.00	40.00	50.00	30	32.98
14	8	5.00	25.00	70.00	23	70.95

Tabla 11-15 Ajuste del modelo para la respuesta dureza

Response: hardness
ANOVA for Mixture Quadratic Model
Analysis of variance table [Partial sum of squares]

Source	Sum of Squares	DF	Mean Square	F Value	Prob > F
Model	279.73	5	55.95	2.37	0.1329
Linear Mixture	*29.13*	*2*	*14.56*	*0.62*	*0.5630*
AB	*72.61*	*1*	*72.61*	*3.08*	*0.1174*
AC	*179.67*	*1*	*179.67*	*7.62*	*0.0247*
BC	*8.26*	*1*	*8.26*	*0.35*	*0.5703*
Residual	188.63	8	23.58		
Lack of Fit	*63.63*	*4*	*15.91*	*0.51*	*0.7354*
Pure Error	*125.00*	*4*	*31.25*		
Cor Total	468.36	13			

Std. Dev.	4.86	R-Squared	0.5973
Mean	24.79	Adj R-Squared	0.3455
C.V.	19.59	Pred R-Squared	−0.3635
PRESS	638.60	Adeq Precision	4.975

Component	Coefficient Estimate	DF	Standard Error	95% CI Low	95% CI High
A-Monomer	23.81	1	3.36	16.07	31.55
B-Crosslinker	16.40	1	7.68	−1.32	34.12
C-Resin	29.45	1	3.36	21.71	37.19
AB	44.42	1	25.31	−13.95	102.80
AC	−44.01	1	15.94	−80.78	−7.25
BC	13.80	1	23.32	−39.97	67.57

Final Equation in Terms of Pseudo Components:

hardness =
+23.81 * A
+16.40 * B
+29.45 * C
+44.42 * A * B
−44.01 * A * C
+13.80 * B * C

do en la figura 11-36 puede generarse utilizando *Design-Expert*. Se supuso que, además de las seis corridas requeridas para ajustar el modelo cuadrático de una mezcla, se harían cuatro corridas diferentes adicionales para verificar la falta de ajuste y que se harían réplicas de cuatro de estas corridas a fin de proporcionar una estimación del error puro. *Design-Expert* utilizó los vértices, los centros en los bordes, el centroide global y las corridas de verificación (los puntos localizados a la mitad entre el centroide y los vértices) como los puntos candidatos.

El diseño con 14 corridas se muestra en la tabla 11-14 junto con las respuestas dureza y sólidos. Los resultados del ajuste de modelos cuadráticos para ambas respuestas se resumen en las tablas 11-15 y 11-16. Observe que los modelos cuadráticos se ajustan muy bien tanto a la respuesta dureza como a la respuesta sólidos. En estas tablas se muestran las ecuaciones ajustadas para ambas respuestas (en términos de los pseudocomponentes). En las figuras 11-37 y 11-38 se muestran las gráficas de contorno de las respuestas.

La figura 11-39 es una gráfica de superposición de las dos superficies de respuesta, donde se muestra el contorno de la dureza Knoop de 25% y el contorno de 30% para los sólidos. La región factible para este producto es el área sin sombrear cerca del centro de la gráfica. Evidentemente, existen varias elecciones

Tabla 11-16 Ajuste del modelo para la respuesta sólidos

Response: solids
ANOVA for Mixture Quadratic Model
Analysis of variance table [Partial sum of squares]

Source	Sum of Squares	DF	Mean Square	F Value	Prob > F
Model	4297.94	5	859.59	25.78	<0.0001
Linear Mixture	*2931.09*	*2*	*1465.66*	*43.95*	*<0.0001*
AB	*211.20*	*1*	*211.20*	*6.33*	*0.0360*
AC	*285.67*	*1*	*285.67*	*8.57*	*0.0191*
BC	*1036.72*	*1*	*1036.72*	*31.09*	*0.0005*
Residual	266.79	8	33.35		
Lack of Fit	*139.92*	*4*	*34.98*	*1.10*	*0.4633*
Pure Error	*126.86*	*4*	*31.72*		
Cor Total	4564.73	13			

Std. Dev.	5.77	R-Squared	0.9416
Mean	33.01	Adj R-Squared	0.9050
C.V.	17.49	Pred R-Squared	0.7827
PRESS	991.86	Adeq Precision	15.075

Component	Coefficient Estimate	DF	Standard Error	95%CI Low	95% CI High
A-Monomer	26.53	1	3.99	17.32	35.74
B-Crosslinker	46.60	1	9.14	25.53	67.68
C-Resin	73.23	1	3.99	64.02	82.43
AB	−75.76	1	30.11	−145.19	−6.34
AC	−55.50	1	18.96	−99.22	−11.77
BC	−154.61	1	27.73	−218.56	−90.67

Final Equation in Terms of Pseudo Components:

solids =
+26.53 * A
+46.60 * B
+73.23 * C
−75.76 * A * B
−55.50 * A * C
−154.61 * B * C

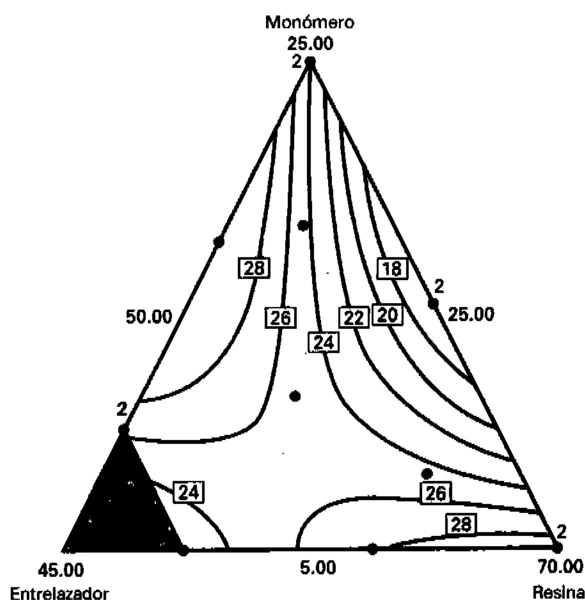


Figura 11-37 Gráfica de contorno de la respuesta dureza Knoop, ejemplo 11-4.

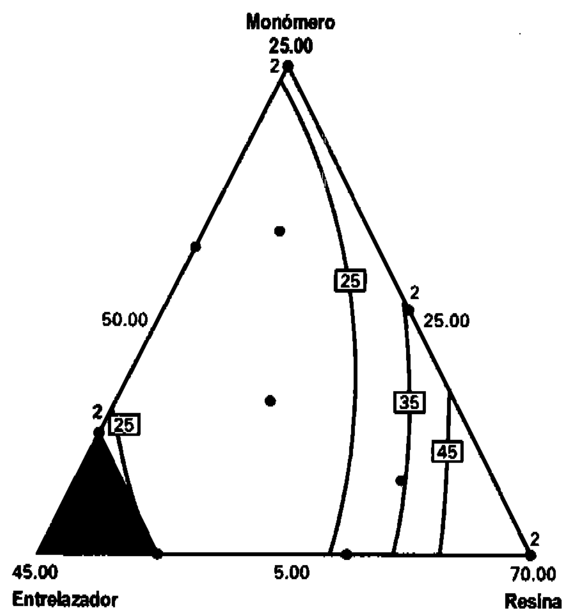


Figura 11-38 Gráfica de contorno de la respuesta sólidos, ejemplo 11-4.

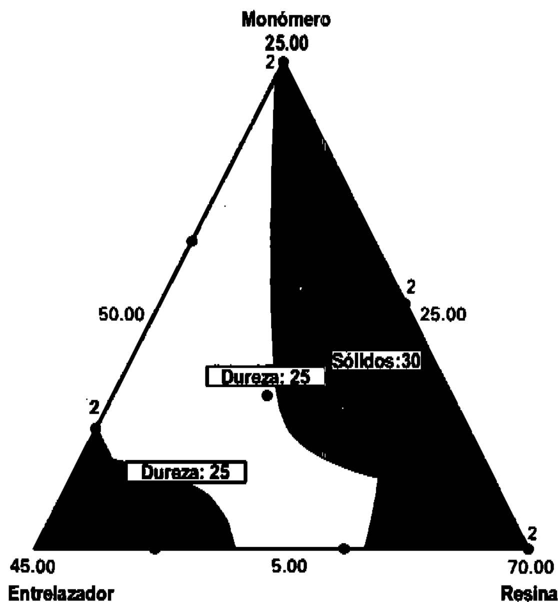


Figura 11-39 Gráfica de contorno de las respuestas dureza Knoop y porcentaje de sólidos, donde se indica la región factible para la formulación de la pintura.

para las proporciones del monómero, el entrelazador y la resina para el recubrimiento total que redundará en un producto que satisfaga los requerimientos de desempeño.

11-6 OPERACIÓN EVOLUTIVA

Personal de investigación y desarrollo aplica con frecuencia la metodología de superficies de respuesta en operaciones de plantas piloto. Cuando se aplica a un proceso de producción a gran escala, suele hacerse una sola vez (o con poca frecuencia), ya que el procedimiento experimental es relativamente minucioso. Sin embargo, las condiciones que fueron óptimas para la planta piloto quizá no lo sean para el proceso a gran escala. La planta piloto puede producir 2 libras de producto por día, mientras que el proceso a gran escala puede generar 2000 libras diarias. Esta "escalación" de la planta piloto al proceso de producción a gran escala da por lo general como resultado la distorsión de las condiciones óptimas. Incluso si la planta a gran escala empieza a operar en el óptimo, con el tiempo se "desvía" de ese punto debido a las variaciones en las materias primas, los cambios ambientales y el personal de operación.

Se requiere un método para el monitoreo y el mejoramiento continuo de un proceso a gran escala cuyo objetivo sea mover las condiciones de operación hacia el óptimo o después de una "desviación". El método no deberá requerir cambios grandes o repentinos de las condiciones de operación que pudieran interrumpir la producción. La **operación evolutiva (EVOP,** por sus siglas en inglés) fue propuesta por Box [12c] como un procedimiento de operación con estas características. Está diseñado como un método de operación rutinaria en la planta que lleva a cabo el personal de manufactura con un mínimo de asistencia del equipo de investigación y desarrollo.

La EVOP consiste en introducir de manera sistemática pequeños cambios en los niveles de las variables de operación bajo consideración. Generalmente se emplea un diseño 2^k para hacer esto. Se supone que los cambios de las variables son lo suficientemente pequeños para que no ocurran perturbaciones serias en el rendimiento, la calidad o la cantidad, pero lo suficientemente grandes para descubrir en última instancia mejoras potenciales en el desempeño del proceso. Se colectan datos de las variables de respuesta de interés en cada punto del diseño 2^k. Cuando se ha hecho una observación en cada punto del diseño, se dice que se ha completado un **ciclo.** Entonces pueden calcularse los efectos y las interacciones de las variables del proceso. Finalmente, después de varios ciclos, el efecto de una o más variables del proceso o sus interacciones sobre la respuesta puede parecer significativo. En este punto se debe tomar una decisión para modificar las condiciones de operación básicas a fin de mejorar la respuesta. Cuando se han detectado las condiciones mejoradas, se dice que se ha terminado una **fase.**

Para probar la significación de las variables e interacciones del proceso, se necesita una estimación del error experimental. Ésta se calcula a partir de los datos del ciclo. Además, el diseño 2^k se centra por lo general en torno a las mejores condiciones de operación actuales. Mediante la comparación de la respuesta en este punto con los 2^k puntos de la porción factorial, es posible verificar la curvatura o **cambio en la media (CIM,** por sus siglas en inglés); es decir, si el proceso en realidad se centra en el máximo, por ejemplo, entonces la respuesta en el centro deberá ser significativamente mayor que las respuestas en los puntos periféricos del diseño 2^k.

En teoría, la EVOP puede aplicarse a *k* variables del proceso. En la práctica es común considerar sólo dos o tres variables. Se presentará un ejemplo del procedimiento para dos variables. Box y Draper [16a] ofrecen un estudio detallado del caso de tres variables, incluyendo las formas y hojas de trabajo necesarias. Myers y Montgomery [85a] revisan la implementación en computadora de la EVOP.

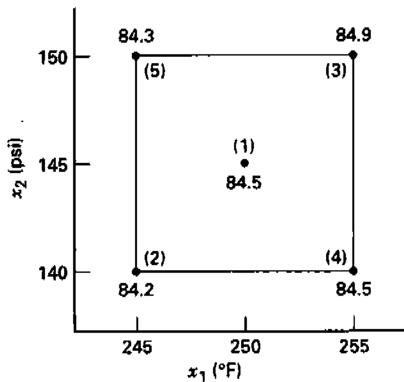


Figura 11-40 Un diseño 2^2 para la EVOP.

EJEMPLO 11-5

Considere un proceso químico cuyo rendimiento es una función de la temperatura (x_1) y la presión (x_2). Las condiciones de operación actuales son x_1 = 250°F y x_2 = 145 psi. El procedimiento EVOP utiliza el diseño 2^2 más el punto central mostrado en la figura 11-40. El ciclo se completa corriendo cada punto del diseño en orden numérico (1, 2, 3, 4, 5). Los rendimientos del primer ciclo se muestran también en la figura 11-40.

Los rendimientos del primer ciclo se anotan en la hoja de cálculo EVOP, la cual se muestra en la tabla 11-17. Al término del primer ciclo no puede hacerse ninguna estimación de la desviación estándar.

Tabla 11-17 Hoja de cálculo EVOP para el ejemplo 11-5, n = 1

5 [•1] 3 2 [ ] 4			Ciclo: n = 1 Respuesta: Rendimiento			Fase: 1 Fecha: 1/11/00
			Cálculo de los promedios			Cálculo de la desviación estándar
Condiciones de operación	(1)	(2)	(3)	(4)	(5)	
(*i*) Suma del ciclo anterior						Suma anterior S =
(*ii*) Promedio del ciclo anterior						Promedio anterior S =
(*iii*) Nuevas observaciones	84.5	84.2	84.9	84.5	84.3	Nueva S = rango × $f_{5,n}$ =
(*iv*) Diferencias [(*ii*) − (*iii*)]						Rango de (*iv*)=
(*v*) Nuevas sumas [(*i*) + (*iii*)]	84.5	84.2	84.9	84.5	84.3	Nueva suma S =
(*vi*) Nuevos promedios [$\bar{y}_i$ = (v)/n]	84.5	84.2	84.9	84.5	84.3	Nuevo promedio $S = \dfrac{\text{Nueva suma } S}{n-1}$

Cálculo de los efectos	Cálculo de los límites de error
Efecto de la temperatura $= \frac{1}{2}(\bar{y}_3 + \bar{y}_4 - \bar{y}_2 - \bar{y}_5) = 0.45$	Para el nuevo promedio $= \dfrac{2}{\sqrt{n}}S =$
Efecto de la presión $= \frac{1}{2}(\bar{y}_3 + \bar{y}_5 - \bar{y}_2 - \bar{y}_4) = 0.25$ Efecto de la interacción $T \times P = \frac{1}{2}(\bar{y}_2 + \bar{y}_3 - \bar{y}_4 - \bar{y}_5) = 0.15$	Para los nuevos efectos $\dfrac{2}{\sqrt{n}}S =$
Efecto del cambio en la media $= \frac{1}{5}(\bar{y}_2 + \bar{y}_3 + \bar{y}_4 + \bar{y}_5 - 4\bar{y}_1) = 0.02$	Para el cambio en la media $\dfrac{1.78}{\sqrt{n}}S =$

Tabla 11-18 Hoja de cálculo EVOP para el ejemplo 11-5, $n = 2$

5 [•1] 3 2 [ ] 4			Ciclo: $n = 2$ Respuesta: Rendimiento			Fase: 1 Fecha: 1/11/00
			Cálculo de los promedios			Cálculo de la desviación estándar
Condiciones de operación	(1)	(2)	(3)	(4)	(5)	
(i) Suma del ciclo anterior	84.5	84.2	84.9	84.5	84.3	Suma anterior S =
(ii) Promedio del ciclo anterior	84.5	84.2	84.9	84.5	84.3	Promedio anterior S =
(iii) Nuevas observaciones	84.9	84.6	85.9	83.5	84.0	Nueva S = rango × $f_{5,n} = 0.60$
(iv) Diferencias [(ii) – (iii)]	–0.4	–0.4	–1.0	+1.0	0.3	Rango de (iv) = 2.0
(v) Nuevas sumas [(i) + (iii)]	169.4	168.8	170.8	168.0	168.3	Nueva suma S = 0.60
(vi) Nuevos promedios [$\bar{y}_i$ = (v)/n]	84.70	84.40	85.40	84.00	84.15	Nuevo promedio $S = \frac{\text{Nueva suma } S}{n-1} = 0.60$

Cálculo de los efectos	Cálculo de los límites de error
Efecto de la temperatura $= \frac{1}{2}(\bar{y}_3 + \bar{y}_4 - \bar{y}_2 - \bar{y}_5) = 0.43$	Para el nuevo promedio $= \frac{2}{\sqrt{n}} S = 0.85$
Efecto de la presión $= \frac{1}{2}(\bar{y}_3 + \bar{y}_5 - \bar{y}_2 - \bar{y}_4) = 0.58$	Para los nuevos efectos $\frac{2}{\sqrt{n}} S = 0.85$
Efecto de la interacción $T \times P = \frac{1}{2}(\bar{y}_2 + \bar{y}_3 - \bar{y}_4 - \bar{y}_5) = 0.83$	
Efecto del cambio en la media $= \frac{1}{5}(\bar{y}_2 + \bar{y}_3 + \bar{y}_4 + \bar{y}_5 - 4\bar{y}_1) =$ –0.17	Para el cambio en la media $\frac{1.78}{\sqrt{n}} S = 0.76$

Los efectos y la interacción de la temperatura y la presión se calculan de la manera usual para un diseño 2^2.

Después se corre un segundo ciclo y los datos del rendimiento se registran en otra hoja de cálculo EVOP, la cual se muestra en la tabla 11-18. Al final del segundo ciclo, el error experimental puede estimarse y las estimaciones de los efectos pueden compararse con límites aproximados de 95% (dos desviaciones estándar). Observe que el rango se refiere al rango de las diferencias del renglón (iv); por lo tanto, el rango es +1.0 – (–1.0) = 2.0. Puesto que ninguno de los efectos de la tabla 11-18 excede sus límites de error, probablemente el verdadero efecto sea cero, y no se contemplan modificaciones en las condiciones de operación.

En la tabla 11-19 se muestran los resultados de un tercer ciclo. Ahora, el efecto de la presión excede su límite de error y el efecto de la temperatura es igual al límite de error. Probablemente ahora se justifique un cambio en las condiciones de operación.

A la luz de los resultados, parece razonable empezar una nueva fase EVOP alrededor del punto (3). Por lo tanto, $x_1 = 225$°F y $x_2 = 150$ psi serían el centro del diseño 2^2 en la segunda fase.

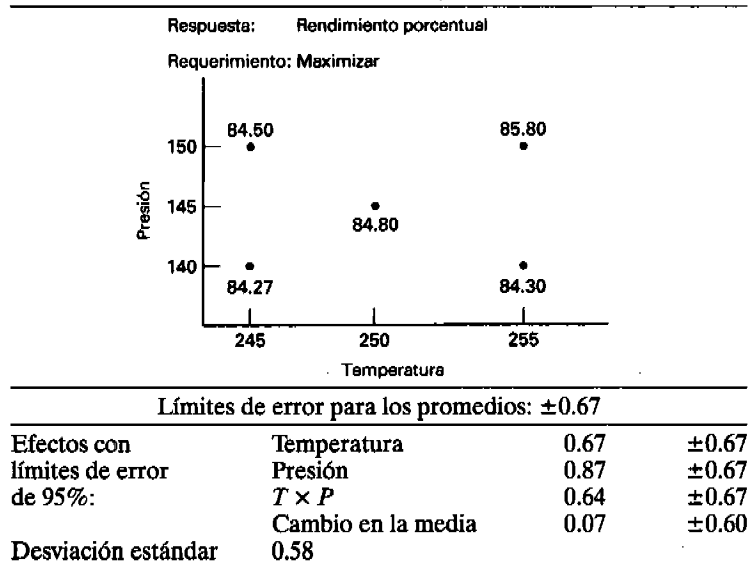
Un aspecto importante de la EVOP es la retroalimentación de información generada por el proceso para operadores y supervisores. Esto se consigue mediante un tablero con información EVOP a la vista de todos. En la tabla 11-20 se muestra el tablero de información para este ejemplo al final del ciclo 3.

Tabla 11-19 Hoja de cálculo EVOP para el ejemplo 11-5, $n = 3$

5 [•1] 3 / 2 [ ] 4			Ciclo: $n = 3$ Respuesta: Rendimiento			Fase: 1 Fecha: 1/11/00
			Cálculo de los promedios			Cálculo de la desviación estándar
Condiciones de operación	(1)	(2)	(3)	(4)	(5)	
(i) Suma del ciclo anterior	169.4	168.8	170.8	168.0	168.3	Suma anterior $S = 0.60$
(ii) Promedio del ciclo anterior	84.70	84.40	85.40	84.00	84.15	Promedio anterior $S = 0.60$
(iii) Nuevas observaciones	85.0	84.0	86.6	84.9	85.2	Nueva S = rango × $f_{5,n} = 0.56$
(iv) Diferencias [(ii) – (iii)]	–0.30	+0.40	–1.20	–0.90	–1.05	Rango de (iv)= 1.60
(v) Nuevas sumas [(i) + (iii)]	254.4	252.8	257.4	252.9	253.5	Nueva suma $S = 1.16$
(vi) Nuevos promedios [$\bar{y}_i$ = (v)/n]	84.80	84.27	85.80	84.30	84.50	Nuevo promedio $S = \dfrac{\text{Nueva suma } S}{n-1} = 0.58$

Cálculo de los efectos	Cálculo de los límites de error
Efecto de la temperatura $= \frac{1}{2}(\bar{y}_3 + \bar{y}_4 - \bar{y}_2 - \bar{y}_5) = 0.67$	Para el nuevo promedio $= \frac{2}{\sqrt{n}} S = 0.67$
Efecto de la presión $= \frac{1}{2}(\bar{y}_3 + \bar{y}_5 - \bar{y}_2 - \bar{y}_4) = 0.87$ Efecto de la interacción $T \times P = \frac{1}{2}(\bar{y}_2 + \bar{y}_3 - \bar{y}_4 - \bar{y}_5) = 0.64$	Para los nuevos efectos $\frac{2}{\sqrt{n}} S = 0.67$
Efecto del cambio en la media $= \frac{1}{5}(\bar{y}_2 + \bar{y}_3 + \bar{y}_4 + \bar{y}_5 - 4\bar{y}_1) = -0.07$	Para el cambio en la media $\frac{1.78}{\sqrt{n}} S = 0.60$

Tabla 11-20 Tablero de información EVOP, ciclo 3



Límites de error para los promedios: ±0.67			
Efectos con límites de error de 95%:	Temperatura	0.67	±0.67
	Presión	0.87	±0.67
	$T \times P$	0.64	±0.67
	Cambio en la media	0.07	±0.60
Desviación estándar	0.58		

Tabla 11-21 Valores de $f_{k,n}$

$n =$	2	3	4	5	6	7	8	9	10
$k = 5$	0.30	0.35	0.37	0.38	0.39	0.40	0.40	0.40	0.41
9	0.24	0.27	0.29	0.30	0.31	0.31	0.31	0.32	0.32
10	0.23	0.26	0.28	0.29	0.30	0.30	0.30	0.31	0.31

La mayoría de las cantidades de la hoja de cálculo EVOP se obtienen directamente del análisis del diseño factorial 2^k. Por ejemplo, la varianza de cualquier efecto, como $\frac{1}{2}(\bar{y}_3 + \bar{y}_5 - \bar{y}_2 - \bar{y}_4)$, es simplemente σ^2/n, donde σ^2 es la varianza de las observaciones (y). Por lo tanto, los límites de error de dos desviaciones estándar (que corresponden a 95%) para cualquier efecto serían $\pm 2\sigma/\sqrt{n}$. La varianza del cambio en la media es

$$V(\text{CIM}) = V\left[\frac{1}{5}(\bar{y}_2 + \bar{y}_3 + \bar{y}_4 + \bar{y}_5 - 4\bar{y}_1)\right]$$
$$= \frac{1}{25}(5\sigma^2_{\bar{y}} + 16\sigma^2_{\bar{y}}) = \left(\frac{20}{25}\right)\frac{\sigma^2}{n}$$

Por lo tanto, los límites de error de dos desviaciones estándar para el CIM son $\pm(2\sqrt{20/25})\sigma/\sqrt{n} = \pm 1.78\sigma/\sqrt{n}$.

La desviación estándar σ se estima por el método del rango. Sea que $y_i(n)$ denote la observación en el punto del diseño i-ésimo en el ciclo n, y que $\bar{y}_i(n)$ denote el promedio correspondiente de $y_i(n)$ después de n ciclos. Las cantidades del renglón (iv) de la hoja de cálculo EVOP son las diferencias $y_i(n) - \bar{y}_i(n-1)$. La varianza de estas diferencias es

$$V[y_i(n) - \bar{y}_i(n-1)] \equiv \sigma^2_D = \sigma^2\left[1 + \frac{1}{(n-1)}\right] = \sigma^2\frac{n}{(n-1)}$$

El rango de las diferencias, por ejemplo R_D, se relaciona con la estimación de la desviación estándar de las diferencias por $\hat{\sigma}_D = R_D/d_2$. El factor d_2 depende del número de observaciones utilizadas para calcular R_D. Entonces $R_D/d_2 = \hat{\sigma}\sqrt{n/(n-1)}$, por lo que puede usarse

$$\hat{\sigma} = \sqrt{\frac{(n-1)}{n}}\frac{R_D}{d_2} = (f_{k,n})R_D \equiv S$$

para estimar la desviación estándar de las observaciones, donde k denota el número de puntos que se utilizaron en el diseño. Para un diseño 2^2 con un punto central se tiene $k = 5$, y para un diseño 2^3 con un punto central se tiene $k = 9$. Los valores de $f_{k,n}$ se dan en la tabla 11-21.

11-7 DISEÑO ROBUSTO

11-7.1 Antecedentes

A lo largo de este libro se ha hecho hincapié en la importancia del uso de experimentos diseñados estadísticamente en el proyecto, desarrollo y mejoramiento de productos y procesos. A partir de la década de 1980, los ingenieros y científicos han adquirido la conciencia creciente de los beneficios del uso de experi-

mentos diseñados y, en consecuencia, ha habido muchas áreas de aplicaciones nuevas. Una de las más importantes de éstas es el **diseño robusto,** donde la atención se centra en uno o más de los siguientes puntos:

1. El diseño de sistemas (productos o procesos) que no sean sensibles a factores ambientales que puedan afectar el desempeño una vez que el sistema se ha desplegado en el campo. Un ejemplo es la formulación de una pintura para exteriores que debe tener gran duración cuando se exponga a una variedad de condiciones climáticas. Puesto que las condiciones climáticas no son del todo predecibles, y ciertamente no son constantes, el responsable de la formulación del producto quiere que éste sea robusto contra un amplio rango de factores de temperatura, humedad y precipitación pluvial que afectan el desgaste y acabado de la pintura.
2. El diseño de productos para que no sean sensibles a la variabilidad transmitida por los componentes del sistema. Un ejemplo es el diseño de un amplificador electrónico para que el voltaje de salida esté tan cerca como sea posible del valor nominal deseado, independientemente de la variabilidad de los parámetros eléctricos de los resistores, transistores y fuentes de poder que son los componentes del aparato.
3. El diseño de procesos para que el producto manufacturado esté tan cerca como sea posible de las especificaciones nominales, aun cuando sea imposible controlar con toda precisión algunas variables del proceso (como la temperatura) o las características de las materias primas.
4. Determinar las condiciones de operación de un proceso para que las características críticas del producto estén tan cerca como sea posible del valor objetivo deseado y la variabilidad en torno a este objetivo se minimice. Ejemplos de este tipo de problema ocurren con frecuencia. Uno de ellos sucede en la manufactura de semiconductores, donde sería deseable que el espesor del óxido de una oblea estuviera lo más cerca posible del espesor objetivo promedio, así como que la variabilidad del espesor a lo largo de la oblea (una medida de uniformidad) fuese lo más pequeña posible.

A principios de la década de 1980, el ingeniero japonés Genichi Taguchi introdujo un enfoque para resolver problemas de este tipo, a los que se hace referencia de manera conjunta como el problema del **diseño paramétrico robusto (RPD,** por sus siglas en inglés) (ver Taguchi y Wu [109] y Taguchi [108a, b]). Su enfoque se basó en la clasificación de las variables de un proceso o producto como **variables de control** (o **controlables**) y **variables de ruido** (o **no controlables**) para después encontrar los ajustes de las variables controlables que minimizan la variabilidad transmitida a la respuesta por las variables no controlables. Se establece el supuesto de que aun cuando los factores de ruido no son controlables en el sistema a gran escala, pueden controlarse para los fines de un experimento. Referirse a la figura 1-1 para una ilustración gráfica de las variables controlables y no controlables en el contexto general de un experimento diseñado.

Taguchi introdujo algunos métodos estadísticos novedosos y ciertas variantes de las técnicas establecidas como parte de este procedimiento RPD. Hizo uso de diseños factoriales altamente fraccionados y otros tipos de diseños fraccionados obtenidos a partir de arreglos ortogonales. Su metodología generó múltiples debates y controversias. Parte de la polémica surgió porque la metodología de Taguchi fue defendida en Occidente inicialmente (y principalmente) por empresarios, y la ciencia estadística subyacente no había sido revisada adecuadamente por los especialistas. Para finales de la década de 1980, los resultados de una revisión muy completa indicaron que aun cuando los conceptos de ingeniería de Taguchi y el objetivo global del RPD tenían bases sólidas, había problemas de fondo con esta estrategia experimental y con los métodos para el análisis de datos. Para detalles específicos de estos temas, ver Box [12d], Box, Bisgaard y Fung [14], Hunter [59a, b], Montgomery [80b], Myers y Montgomery [85a] y Pignatiello y Ramberg [94]. Muchas de estas preocupaciones se encuentran resumidas también en el amplio panel de

Tabla 11-22 Diseño paramétrico con arreglos tanto interior como exterior

						b) Arreglo exterior							
					E	1	1	1	1	2	2	2	2
					F	1	1	2	2	1	1	2	2
					G	1	2	1	2	1	2	1	2
	a) Arreglo interior												
Corrida	*A*	*B*	*C*	*D*									
1	1	1	1	1		15.6	9.5	16.9	19.9	19.6	19.6	20.0	19.1
2	1	2	2	2		15.0	16.2	19.4	19.2	19.7	19.8	24.2	21.9
3	1	3	3	3		16.3	16.7	19.1	15.6	22.6	18.2	23.3	20.4
4	2	1	2	3		18.3	17.4	18.9	18.6	21.0	18.9	23.2	24.7
5	2	2	3	1		19.7	18.6	19.4	25.1	25.6	21.4	27.5	25.3
6	2	3	1	2		16.2	16.3	20.0	19.8	14.7	19.6	22.5	24.7
7	3	1	3	2		16.4	19.1	18.4	23.6	16.8	18.6	24.3	21.6
8	3	2	1	3		14.2	15.6	15.1	16.8	17.8	19.6	23.2	24.2
9	3	3	2	1		16.1	19.9	19.3	17.3	23.1	22.7	22.6	28.6

discusión publicado en *Technometrics* (ver Nair, *et al.* [86]). En el material suplementario del texto de este capítulo también se comentan e ilustran muchos de los problemas implícitos en los métodos técnicos de Taguchi.

La metodología de Taguchi para el problema RPD gira en torno al uso de un diseño ortogonal para los factores controlables, el cual se "cruza" con un diseño ortogonal separado para los factores de ruido. En la tabla 11-22 se presenta un ejemplo de Byrne y Taguchi [23] que trata del desarrollo de un método para ensamblar un conector elastométrico en un tubo de nylon que produciría la fuerza de separación requerida. Hay cuatro factores controlables, cada uno con tres niveles (A = interferencia, B = espesor de la pared del conector, C = profundidad de inserción y D = porcentaje de adhesivo), y tres factores de ruido o no controlables (E = tiempo de acondicionamiento, F = temperatura de acondicionamiento y G = humedad relativa del acondicionamiento). El panel *a* de la tabla 11-22 contiene el diseño para los factores controlables. Observe que se trata de un diseño factorial fraccionado de tres niveles; específicamente, es un diseño 3^{4-2}. Taguchi lo llama el **diseño de arreglo interior.** El panel *b* de la tabla 11-22 contiene un diseño 2^3 para los factores de ruido, al que Taguchi llama el **diseño de arreglo exterior.** Entonces se realiza cada corrida del arreglo interior para todas las combinaciones de tratamientos del arreglo exterior, produciéndose las 72 observaciones de la fuerza de separación que se muestran en la tabla. A este tipo de diseño se le llama **diseño de arreglo cruzado.**

Taguchi sugirió que los datos de un experimento de arreglo cruzado se resumieran con dos estadísticos: el promedio de cada observación en el arreglo interior para todas las corridas del arreglo exterior y un resumen de estadísticas que intentaba combinar información acerca de la media y la varianza, llamado **relación señal a ruido.** Las relaciones señal a ruido se definen a propósito para que un valor máximo de la relación minimice la variabilidad transmitida por las variables de ruido. Entonces se lleva a cabo un análisis para determinar cuáles son los ajustes de los factores controlables que dan como resultado 1) una media tan próxima como sea posible al objetivo deseado y 2) un valor máximo de la relación señal a ruido.

El examen de la tabla 11-22 revela un problema importante con la estrategia de diseño de Taguchi; a saber, el enfoque del arreglo cruzado llevará a un experimento muy grande. En el ejemplo tratado aquí sólo hay siete factores, pero el diseño tiene 72 corridas. Además, el diseño de arreglo interno es un diseño 3^{4-2} de resolución III (ver el capítulo 9 para un estudio de este diseño), por lo que a pesar del gran número de corridas, no es posible obtener *ninguna* información acerca de las interacciones entre las variables controlables. De hecho, incluso la información acerca de los efectos principales está potencialmente corrompida, ya que los efectos principales tienen estrechas relaciones de alias con las interacciones de dos factores. Ocurre también que las relaciones señal a ruido de Taguchi son problemáticas; al maximizarse la relación no se minimiza necesariamente la variabilidad. Referirse al material complementario del texto para mayores detalles.

Un punto importante acerca del diseño de arreglo cruzado es que *sí* proporciona información acerca de las interacciones factor controlable × factor de ruido. Estas interacciones son cruciales para la solución de un problema RPD. Por ejemplo, considere las gráficas de las interacciones de dos factores de la figura 11-41, donde x es el factor controlable y z el factor de ruido. En la figura 11-41*a* no hay ninguna interacción $x \times z$; por lo tanto, no hay ningún valor de la variable controlable x que afecte la variabilidad transmitida a la respuesta por la variabilidad en z. Sin embargo, en la figura 11-41*b* hay una fuerte interacción $x \times z$. Observe que cuando x se pone en el nivel bajo, hay mucho menos variabilidad en la variable de respuesta que cuando x está en el nivel alto. Por lo tanto, a menos que haya como mínimo una interacción factor controlable × factor de ruido, no hay ningún problema de diseño robusto. Como se verá en la siguiente sección, enfocarse en la identificación y el modelado de estas interacciones es una de las claves de un enfoque más eficiente y eficaz del RPD.

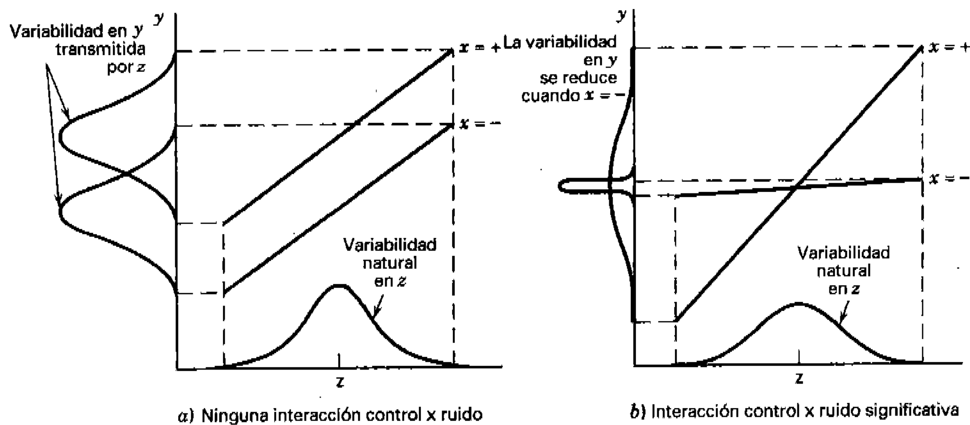


Figura 11-41 El papel de la interacción control × ruido en un diseño robusto.

11-7.2 El enfoque de la superficie de respuesta para el diseño robusto

Como se señaló en la sección anterior, las interacciones entre los factores controlables y los de ruido son la clave en un problema de diseño robusto. Por lo tanto, es lógico usar un **modelo** de respuesta que incluya tanto a los factores controlables como a los factores de ruido y sus interacciones. Para ilustrar, suponga que se tienen dos factores controlables x_1 y x_2 y un solo factor de ruido z_1. Se supone que tanto los factores controlables como el de ruido se expresan como las variables codificadas usuales (es decir, tienen su centro en cero y tienen límites inferior y superior $\pm a$). Si quiere considerarse un modelo de primer orden que incluya las variables controlables, un modelo lógico es

$$y=\beta_0+\beta_1x_1+\beta_2x_2+\beta_{12}x_1x_2+\gamma_1z_1+\delta_{11}x_1z_1+\delta_{21}x_2z_1+\varepsilon \qquad (11\text{-}28)$$

Observe que este modelo incluye los efectos principales de ambos factores controlables, su interacción, el efecto principal de la variable de ruido y las dos interacciones entre las variables controlables y la de ruido. A este tipo de modelo, el cual incorpora a las variables controlables y las de ruido, suele llamársele **modelo de respuesta** o **de reacción.** Excepto cuando al menos uno de los coeficientes de regresión δ_{11} y δ_{21} sea diferente de cero, no habrá ningún problema de diseño robusto.

Una ventaja importante del enfoque del modelo de respuesta es que tanto los factores controlables como los factores de ruido pueden colocarse en un solo diseño experimental; es decir, puede evitarse la estructura de los arreglos interior y exterior del enfoque de Taguchi. Al diseño que contiene tanto los factores controlables como los de ruido suele llamársele **diseño de arreglo combinado.**

Como se señaló anteriormente, se supone que las variables de ruido son aleatorias, aun cuando son controlables para los fines de un experimento. Específicamente, se supone que las variables de ruido están expresadas en unidades codificadas, que tienen valor esperado cero, varianza σ_z^2, y que si hay varias variables de ruido, tienen covarianzas cero. Bajo estos supuestos es sencillo encontrar un modelo para la respuesta media tomando el valor esperado de y en la ecuación 11-28. Se obtiene así

$$E_z(y)=\beta_0+\beta_1x_1+\beta_2x_2+\beta_{12}x_1x_2$$

donde el subíndice z del operador expectativa es un recordatorio para tomar el valor esperado con respecto a *ambas* variables aleatorias de la ecuación 11-28, z_1 y ε. Para encontrar un modelo de la varianza de la

respuesta y se usa el **enfoque de la transmisión del error.** Primero, el modelo de respuesta de la ecuación 11-28 se expande en una serie de Taylor de primer orden alrededor de $z_1 = 0$. Se obtiene así

$$\begin{aligned} y &\cong y_{z=0} + \frac{dy}{dz_1}(z_1 - 0) + R + \varepsilon \\ &\cong \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 \\ &\quad + (\gamma_1 + \delta_{11} x_1 + \delta_{21} x_2) z_1 + R + \varepsilon \end{aligned}$$

donde R es el término del residuo de la serie de Taylor. Como es común en la práctica, se ignorará el término del residuo. Ahora puede obtenerse la varianza de y aplicando el operador varianza en esta última expresión (sin R). El modelo para la varianza resultante es

$$V_z(y) = \sigma_z^2(\gamma_1 + \delta_{11}x_1 + \delta_{21}x_2)^2 + \sigma^2$$

De nueva cuenta se ha usado el subíndice z en el operador varianza como recordatorio de que *tanto* z_1 *como* ε son variables aleatorias.

Se han derivado modelos simples para la media y la varianza de la variable de respuesta de interés. Observe lo siguiente:

1. Los modelos de la media y la varianza incluyen **únicamente las variables controlables.** Esto significa que es potencialmente posible fijar las variables controlables para alcanzar un valor objetivo de la media y minimizar la variabilidad transmitida por la variable de ruido.
2. Aun cuando en el modelo de la varianza intervienen sólo las variables controlables, incluye asimismo los *coeficientes de regresión de la interacción* entre las variables controlables y la de ruido. Es así como la variable de ruido influye en la respuesta.
3. El modelo de la varianza es una **función cuadrática** de las variables controlables.
4. El modelo de la varianza (dejando de lado σ^2) es sólo el cuadrado de la **pendiente** del modelo de respuesta ajustado en la dirección de la variable de ruido.

Para dar un uso operacional a estos modelos sería necesario:

1. Realizar un experimento y ajustar un modelo de respuesta apropiado, tal como la ecuación 11-28.
2. Sustituir los coeficientes de regresión desconocidos en los modelos de la media y la varianza con sus estimaciones de mínimos cuadrados del modelo de la respuesta o de reacción, y sustituir la σ^2 del modelo de la varianza con el cuadrado medio de los residuales que se encontró cuando se ajustó el modelo de respuesta.
3. Optimizar los modelos de la media y la varianza utilizando los métodos estándares de optimización de respuestas múltiples revisadas en la sección 11-3.4.

Es muy sencillo generalizar estos resultados. Suponga que hay k variables controlables y r variables de ruido. El modelo de respuesta general que incluye estas variables se escribirá como

$$y(\mathbf{x}, \mathbf{z}) = f(\mathbf{x}) + h(\mathbf{x}, \mathbf{z}) + \varepsilon \qquad (11\text{-}29)$$

donde $f(\mathbf{x})$ es la porción del modelo que incluye sólo las variables controlables y $h(\mathbf{x}, \mathbf{z})$ son los términos que incluyen los efectos principales de los factores de ruido y las interacciones entre los factores controlables y los de ruido. De manera típica, la estructura de $h(\mathbf{x}, \mathbf{z})$ es

$$h(\mathbf{x}, \mathbf{z}) = \sum_{i=1}^{r} \gamma_i z_i + \sum_{i=1}^{k}\sum_{j=1}^{r} \delta_{ij} x_i z_j$$

La estructura de $f(\mathbf{x})$ dependerá de cuál sea el tipo de modelo que el experimentador considere apropiado para las variables controlables. Las elecciones lógicas son el modelo de primer orden con interacción y el modelo de segundo orden. Si se supone que las variables de ruido tienen media cero, varianza σ_z^2 y covarianzas cero, y que las variables de ruido y los errores aleatorios ε tienen covarianzas cero, entonces el modelo de la media para la respuesta es

$$E_z[y(\mathbf{x}, \mathbf{z})] = f(\mathbf{x}) \tag{11-30}$$

y el modelo de la varianza para la respuesta es

$$V_z[y(\mathbf{x}, \mathbf{z})] = \sigma_z^2 \sum_{i=1}^{r} \left[\frac{\partial y(\mathbf{x}, \mathbf{z})}{\partial z_i}\right]^2 + \sigma^2 \tag{11-31}$$

Myers y Montgomery [85a] presentan una forma un tanto más general de la ecuación 11-31 basada en la aplicación directa de un operador de varianza condicional al modelo de respuesta.

EJEMPLO 11-6 ··

Para ilustrar el procedimiento anterior, considere nuevamente el ejemplo 6-2 en el que se estudiaron cuatro factores en un diseño factorial 2^4 para investigar su efecto sobre la rapidez de filtración de un producto químico. Se supondrá que el factor A, la temperatura, es difícil de controlar en el proceso a gran escala, pero que puede controlarse durante el experimento (el cual se llevó a cabo en una planta piloto). Los otros tres factores, la presión (B), la concentración (C) y la velocidad de agitación (D), son fáciles de controlar. Por lo tanto, el factor de ruido z_1 es la temperatura, y las variables controlables x_1, x_2 y x_3 son la presión, la concentración y la velocidad de agitación, respectivamente. Puesto que tanto los factores controlables como el factor de ruido están en el mismo diseño, el diseño factorial 2^4 utilizado en este experimento es un ejemplo de un **diseño de arreglo combinado.**

Utilizando los resultados del ejemplo 6-2, el modelo de respuesta es

$$\begin{aligned}\hat{y}(\mathbf{x}, z_1) &= 70.06 + \left(\frac{21.625}{2}\right) z_1 + \left(\frac{9.875}{2}\right) x_2 + \left(\frac{14.625}{2}\right) x_3 \\ &\quad - \left(\frac{18.125}{2}\right) x_2 z_1 + \left(\frac{16.625}{2}\right) x_3 z_1 \\ &= 70.06 + 10.81 z_1 + 4.94 x_2 + 7.31 x_3 - 9.06 x_2 z_1 + 8.31 x_3 z_1\end{aligned}$$

Utilizando las ecuaciones 11-30 y 11-31, se encuentra que los modelos de la media y la varianza son

$$E_z[y(\mathbf{x}, z_1)] = 70.06 + 4.94 x_2 + 7.31 x_3$$

y

$$\begin{aligned}V_z[y(\mathbf{x}, z_1)] &= \sigma_z^2 (10.81 - 9.06 x_2 + 8.31 x_3)^2 + \sigma^2 \\ &= \sigma_z^2 (116.91 + 82.08 x_2^2 + 69.06 x_3^2 - 195.88 x_2 \\ &\quad + 179.66 x_3 - 150.58 x_2 x_3) + \sigma^2\end{aligned}$$

respectivamente. Suponga ahora que los niveles bajo y alto de la variable de ruido, temperatura, se corrieron a una desviación estándar a ambos lados de su valor típico o promedio, de tal modo que $\sigma_z^2 = 1$, y que se usa $\hat{\sigma}^2 = 19.51$ (éste es el cuadrado medio de los residuales obtenido al ajustar el modelo de respuesta). Por lo tanto, el modelo de la varianza queda como

$$V_z[y(\mathbf{x}, z_1)] = 136.42 - 195.88 x_2 + 179.66 x_3 - 150.58 x_2 x_3 + 82.08 x_2^2 + 69.06 x_3^2$$

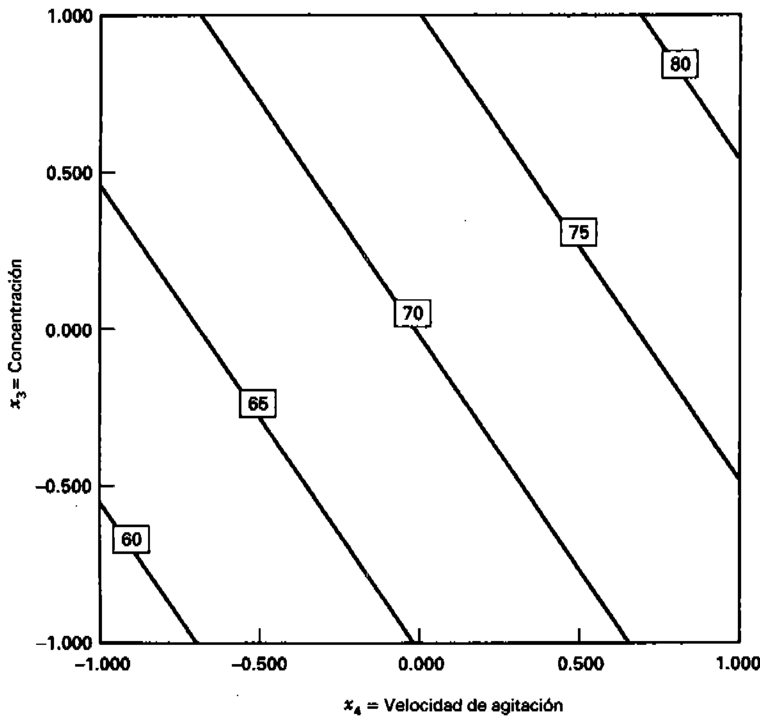
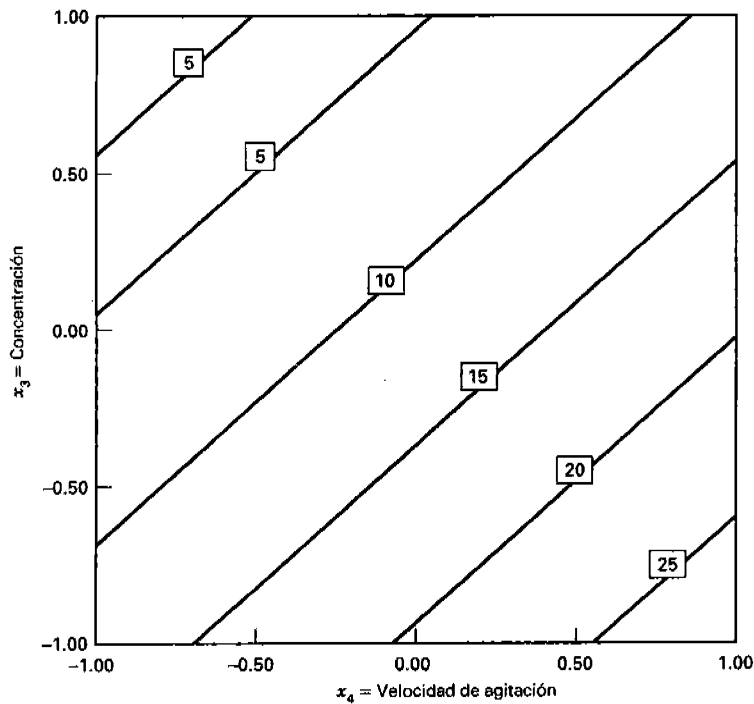


Figura 11-42 Contornos del índice de filtración medio constante, ejemplo 11-6, con x_1 = temperatura = 0.

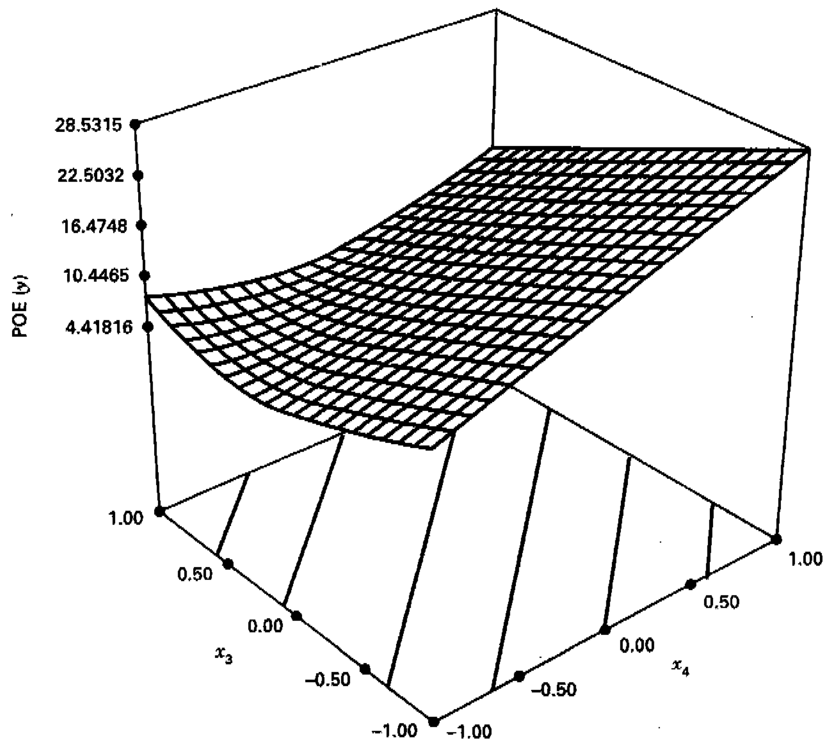
En la figura 11-42 se presenta la gráfica de contorno del paquete de software *Design-Expert* de los contornos de respuesta del modelo de la media. Para construir esta gráfica se fijó el factor de ruido (temperatura) en cero y el factor controlable no significativo (presión) también en cero. Observe que la rapidez de filtración promedio se incrementa cuando tanto la concentración como la velocidad de agitación se incrementan. *Design-Expert* constituirá también de manera automática gráficas de la **raíz cuadrada** de los contornos de la varianza, que denomina **propagación del error** (o **POE,** por sus siglas en inglés). Evidentemente, la POE no es sino la desviación estándar de la variabilidad que se transmite a la respuesta como una función de las variables controlables. En la figura 11-43 se muestra la gráfica de contorno y la gráfica de superficie de respuesta tridimensional de la POE, obtenida con *Design-Expert* (en esta gráfica la variable de ruido se mantiene constante en cero, como se explicó anteriormente).

Suponga que el experimentador quiere mantener una rapidez de filtración promedio de cerca de 75 y minimizar la variabilidad alrededor de este valor. En la figura 11-44 se muestra una gráfica de superposición de los contornos de la rapidez de filtración media y la POE como una función de la concentración y la velocidad de agitación, las variables controlables significativas. Para conseguir los objetivos deseados será necesario mantener la concentración en el nivel alto y la velocidad de agitación muy cerca del nivel intermedio.

El ejemplo 11-6 ilustra el uso de un modelo de primer orden con interacción como el modelo para los factores controlables, $f(\mathbf{x})$. Se presenta ahora un ejemplo adaptado de Montgomery [80b] que incluye un modelo de segundo orden.



a) Gráfica de contorno



b) Gráfica de superficie de respuesta

Figura 11-43 Gráfica de contorno y superficie de respuesta de la propagación del error (POE) en el ejemplo 11-6, con x_1 = temperatura = 0.

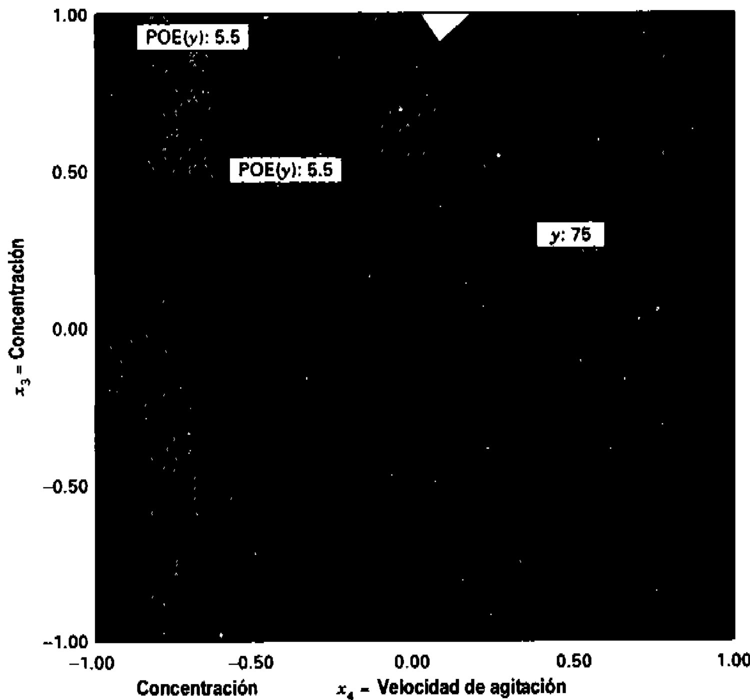


Figura 11-44 Gráfica de superposición de los contornos de la media y la POE del índice de filtración, ejemplo 11-6, con x_1 = temperatura = 0.

EJEMPLO 11-7 ··

En una fábrica de semiconductores se realizó un experimento que incluyó dos variables controlables y tres variables de ruido. En la tabla 11-23 se muestra el diseño de arreglo combinado utilizado por los experimentadores. El diseño es una variante de 23 corridas de un diseño central compuesto que se creó empezando con un DCC estándar para cinco factores (la porción del cubo es un diseño 2^{5-1}) y eliminando las corridas axiales asociadas con las tres variables de ruido. Este diseño soportará un modelo de respuesta que tiene un modelo de segundo orden en las variables controlables, los efectos principales de las tres variables de ruido y las interacciones entre los factores controlables y los de ruido. El modelo de respuesta ajustado es

$$\begin{aligned}\hat{y}(\mathbf{x}, \mathbf{z}) = {} & 30.37 - 2.92x_1 - 4.13x_2 + 2.60x_1^2 + 2.18x_2^2 + 2.87x_1x_2 \\ & + 2.73z_1 - 2.33z_2 + 2.33z_3 - 0.27x_1z_1 + 0.89x_1z_2 + 2.58x_1z_3 \\ & + 2.01x_2z_1 - 1.43x_2z_2 + 1.56x_2z_3\end{aligned}$$

Los modelos de la media y la varianza son

$$E_z[y(\mathbf{x}, \mathbf{z})] = 30.37 - 2.92x_1 - 4.13x_2 + 2.60x_1^2 + 2.18x_2^2 + 2.87x_1x_2$$

y

$$V_z[y(\mathbf{x}, \mathbf{z})] = 19.26 + 3.20x_1 + 12.45x_2 + 7.52x_1^2 + 8.52x_2^2 + 2.21x_1x_2$$

donde se han sustituido las estimaciones de los parámetros del modelo de respuesta ajustado en las ecuaciones de los modelos de la media y la varianza y, como en el ejemplo anterior, se supone que $\sigma_z^2 = 1$. En las figuras 11-45 y 11-46 se presentan las gráficas (de *Design-Expert*) de contorno de la media y la POE del proceso (recuerde que la POE es la raíz cuadrada de la varianza de la superficie de respuesta) generadas a partir de estos modelos.

Tabla 11-23 Experimento de arreglo combinado con dos variables controlables y tres variables de ruido, ejemplo 11-7

Número de corrida	x_1	x_2	z_1	z_2	z_3	y
1	−1.00	−1.00	−1.00	−1.00	1.00	44.2
2	1.00	−1.00	−1.00	−1.00	−1.00	30.0
3	−1.00	1.00	−1.00	−1.00	−1.00	30.0
4	1.00	1.00	−1.00	−1.00	1.00	35.4
5	−1.00	−1.00	1.00	−1.00	−1.00	49.8
6	1.00	−1.00	1.00	−1.00	1.00	36.3
7	−1.00	1.00	1.00	−1.00	1.00	41.3
8	1.00	1.00	1.00	−1.00	−1.00	31.4
9	−1.00	−1.00	−1.00	1.00	−1.00	43.5
10	1.00	−1.00	−1.00	1.00	1.00	36.1
11	−1.00	1.00	−1.00	1.00	1.00	22.7
12	1.00	1.00	−1.00	1.00	−1.00	16.0
13	−1.00	−1.00	1.00	1.00	1.00	43.2
14	1.00	−1.00	1.00	1.00	−1.00	30.3
15	−1.00	1.00	1.00	1.00	−1.00	30.1
16	1.00	1.00	1.00	1.00	1.00	39.2
17	−2.00	0.00	0.00	0.00	0.00	46.1
18	2.00	0.00	0.00	0.00	0.00	36.1
19	0.00	−2.00	0.00	0.00	0.00	47.4
20	0.00	2.00	0.00	0.00	0.00	31.5
21	0.00	0.00	0.00	0.00	0.00	30.8
22	0.00	0.00	0.00	0.00	0.00	30.7
23	0.00	0.00	0.00	0.00	0.00	31.0

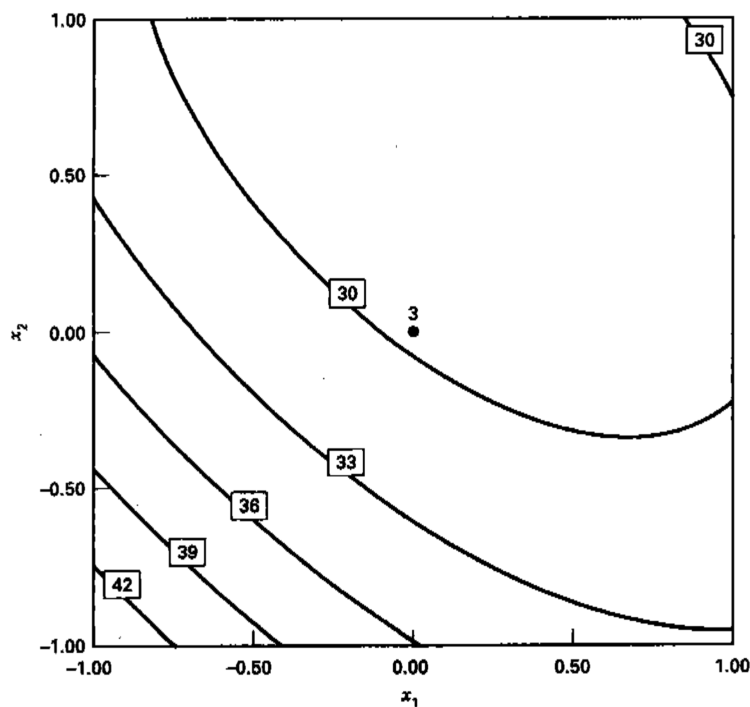


Figura 11-45 Gráfica de contorno del modelo de la media, ejemplo 11-7.

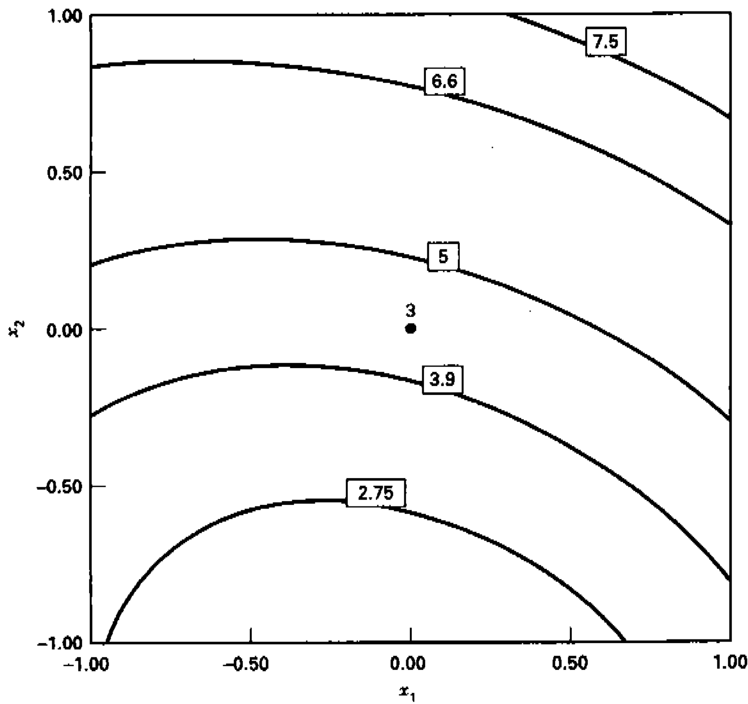


Figura 11-46 Gráfica de contorno de la POE, ejemplo 11-7.

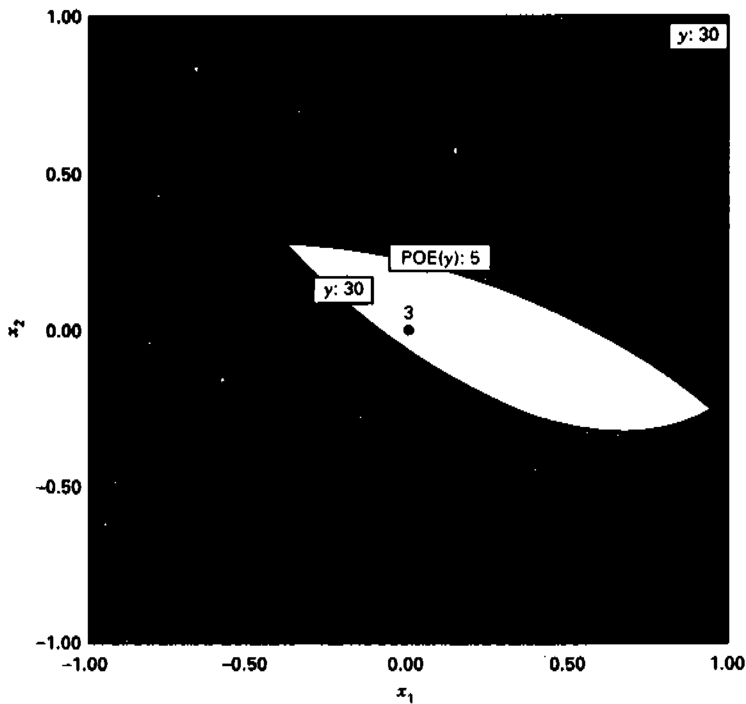


Figura 11-47 Superposición de los contornos de la media y la POE para el ejemplo 11-7, con la región en blanco indicando condiciones de operación satisfactorias para la media y la varianza.

En este problema es deseable mantener la media del proceso abajo de 30. Al inspeccionar las figuras 11-45 y 11-46, es claro que se necesitará hacer un ajuste si se quiere hacer pequeña la varianza del proceso. Puesto que sólo hay dos variables controlables, una forma lógica de llegar a este arreglo es superponer los contornos de la respuesta media constante y la varianza constante, como se muestra en la figura 11-47. Esta gráfica muestra los contornos para los que la media del proceso es menor o igual que 30 y la desviación estándar del proceso es menor o igual que 5. La región delimitada por estos contornos representaría una zona de operación típica de respuesta media baja y varianza del proceso baja.

11-8 PROBLEMAS

11-1. En una planta química se produce oxígeno licuando aire y separándolo por destilación fraccionada en sus gases componentes. La pureza del oxígeno es una función de la temperatura del condensador principal y de la relación de la presión entre las columnas superior e inferior. Las condiciones de operación actuales son temperatura (ξ_1) = –220°C y la relación de la presión (ξ_2) = 1.2. Utilizando los datos siguientes, encontrar la trayectoria del ascenso más pronunciado:

Temperatura (ξ_1)	Índice de la presión (ξ_2)	Pureza
–225	1.1	82.8
–225	1.3	83.5
–215	1.1	84.7
–215	1.3	85.0
–220	1.2	84.1
–220	1.2	84.5
–220	1.2	83.9
–220	1.2	84.3

11-2. Un ingeniero industrial ha desarrollado un modelo de simulación por computadora para un sistema de inventario de dos artículos. Las variables de decisión son la cantidad del pedido y el punto de reorden de cada artículo. La respuesta que debe minimizarse es el costo total del inventario. El modelo de simulación se usa para producir los datos que se muestran en la tabla siguiente. Identificar el diseño experimental. Encontrar la trayectoria del descenso más pronunciado.

Artículo 1		Artículo 2		
Cantidad del pedido (ξ_1)	Punto de reorden (ξ_2)	Cantidad del pedido (ξ_3)	Punto de reorden (ξ_4)	Costo total
100	25	250	40	625
140	45	250	40	670
140	25	300	40	663
140	25	250	80	654
100	45	300	40	648
100	45	250	80	634
100	25	300	80	692
140	45	300	80	686
120	35	275	60	680
120	35	275	60	674
120	35	275	60	681

11-3. Verificar que el siguiente diseño es símplex. Ajustar el modelo de primer orden y encontrar la trayectoria del ascenso más pronunciado.

x_1	x_2	x_3	y
0	$\sqrt{2}$	–1	18.5
$-\sqrt{2}$	0	1	19.8
0	$-\sqrt{2}$	–1	17.4
$\sqrt{2}$	0	1	22.5

11-4. Para el modelo de primer orden

$$\hat{y} = 60 + 1.5x_1 - 0.8x_2 + 2.0x_3$$

encontrar la trayectoria del ascenso más pronunciado. Las variables están codificadas como $-1 \leq x_i \leq 1$.

11-5. La región de experimentación de tres factores son el tiempo ($40 \leq T_1 \leq 80$ min), la temperatura ($200 \leq T_2 \leq 300$°C) y la presión ($20 \leq P \leq 50$ psig). Se ha ajustado un modelo de primer orden en variables codificadas a los datos del rendimiento de un diseño 2^3. El modelo es

$$\hat{y} = 30 + 5x_1 + 2.5x_2 + 3.5x_3$$

¿El punto $T_1 = 85$, $T_2 = 325$, $P = 60$ está en la trayectoria del ascenso más pronunciado?

11-6. La región de experimentación de dos factores son la temperatura ($100 \leq T \leq 300$°F) y la velocidad de alimentación del catalizador ($10 \leq C \leq 30$ lb/pulg). Un modelo de primer orden con las variables codificadas usuales ± 1 se ha ajustado a la respuesta peso molecular, obteniéndose el modelo siguiente:

$$\hat{y} = 2000 + 125x_1 + 40x_2$$

a) Encontrar la trayectoria del ascenso más pronunciado.

b) Se desea mover a una región donde los pesos moleculares rebasen 2500. Con base en la información que se tiene por la experimentación en esta región, ¿aproximadamente cuántos pasos en la trayectoria del ascenso más pronunciado se necesitan para moverse a la región de interés?

11-7. La trayectoria del ascenso más pronunciado suele calcularse suponiendo que el modelo es en realidad de primer orden; es decir, que no hay interacción. Sin embargo, incluso si hay interacción, el ascenso más pronunciado que se determina ignorando la interacción seguirá produciendo por lo general buenos resultados. Para ilustrar, suponga que se ha ajustado el modelo

$$\hat{y} = 20 + 5x_1 - 8x_2 + 3x_1x_2$$

utilizando variables codificadas ($-1 \leq x_i \leq +1$).

a) Trazar la trayectoria del ascenso más pronunciado que se obtendría si se ignorara la interacción.

b) Trazar la trayectoria del ascenso más pronunciado que se obtendría incluyendo la interacción en el modelo. Compararla con la trayectoria que se encontró en el inciso *a*.

11-8. Los datos que se muestran en la siguiente tabla se recolectaron en un experimento para optimizar el crecimiento de un cristal como una función de tres variables x_1, x_2 y x_3. Son deseables los valores grandes de y (rendimiento en gramos). Ajustar un modelo de segundo orden y analizar la superficie ajustada. ¿Bajo qué conjunto de condiciones se alcanza el crecimiento máximo?

x_1	x_2	x_3	y
–1	–1	–1	66
–1	–1	1	70
–1	1	–1	78
–1	1	1	60
1	–1	–1	80
1	–1	1	70
1	1	–1	100
1	1	1	75
–1.682	0	0	100
1.682	0	0	80
0	–1.682	0	68
0	1.682	0	63
0	0	–1.682	65
0	0	1.682	82
0	0	0	113
0	0	0	100
0	0	0	118
0	0	0	88
0	0	0	100
0	0	0	85

11-9. Un ingeniero químico recolectó los siguientes datos. La respuesta y es el tiempo de filtración, x_1 es la temperatura y x_2 es la presión. Ajustar un modelo de segundo orden.

x_1	x_2	y
–1	–1	54
–1	1	45
1	–1	32
1	1	47
–1.414	0	50
1.414	0	53
0	–1.414	47
0	1.414	51
0	0	41
0	0	39
0	0	44
0	0	42
0	0	40

a) ¿Qué condiciones de operación se recomendarían si el objetivo es minimizar el tiempo de filtración?

b) ¿Qué condiciones de operación se recomendarían si el objetivo es operar el proceso con una velocidad de filtración media muy próxima a 46?

11-10. El diseño hexagonal que se presenta a continuación se usa en un experimento que tiene como objetivo ajustar un modelo de segundo orden:

x_1	x_2	y
1	0	68
0.5	$\sqrt{0.75}$	74
−0.5	$\sqrt{0.75}$	65
−1	0	60
−0.5	$-\sqrt{0.75}$	63
0.5	$-\sqrt{0.75}$	70
0	0	58
0	0	60
0	0	57
0	0	55
0	0	69

a) Ajustar el modelo de segundo orden.
b) Efectuar el análisis canónico. ¿Qué tipo de superficie se ha encontrado?
c) ¿Qué condiciones de operación para x_1 y x_2 llevan al punto estacionario?
d) ¿Dónde se correría este proceso si el objetivo es obtener una respuesta que esté tan cerca de 65 como sea posible?

11-11. Un experimentador corrió un diseño de Box-Behnken y obtuvo los siguientes resultados, donde la variable de respuesta es la viscosidad de un polímero:

Nivel	Temperatura	Velocidad de agitación	Presión	x_1	x_2	x_3
Alto	200	10.0	25	+1	+1	+1
Intermedio	175	7.5	20	0	0	0
Bajo	150	5.0	15	−1	−1	−1

Corrida	x_1	x_2	x_3	y_1
1	−1	−1	0	535
2	+1	−1	0	580
3	−1	+1	0	596
4	+1	+1	0	563
5	−1	0	−1	645
6	+1	0	−1	458
7	−1	0	+1	350
8	+1	0	+1	600
9	0	−1	−1	595
10	0	+1	−1	648
11	0	−1	+1	532
12	0	+1	+1	656
13	0	0	0	653
14	0	0	0	599
15	0	0	0	620

a) Ajustar el modelo de segundo orden.
b) Efectuar el análisis canónico. ¿Qué tipo de superficie se ha encontrado?

c) ¿Qué condiciones de operación para x_1, x_2 y x_3 llevan al punto estacionario?

d) ¿Qué condiciones de operación se recomendarían si es importante obtener una viscosidad que esté tan cerca de 600 como sea posible?

11-12. Considere el diseño central compuesto de tres variables que se muestra a continuación. Analizar los datos y sacar conclusiones, suponiendo que se quiere maximizar la conversión (y_1) con la actividad (y_2) entre 55 y 60.

Corrida	Tiempo (min)	Temperatura (°C)	Catalizador (%)	Conversión (%) y_1	Actividad y_2
1	–1.000	–1.000	–1.000	74.00	53.20
2	1.000	–1.000	–1.000	51.00	62.90
3	–1.000	1.000	–1.000	88.00	53.40
4	1.000	1.000	–1.000	70.00	62.60
5	–1.000	–1.000	1.000	71.00	57.30
6	1.000	–1.000	1.000	90.00	67.90
7	–1.000	1.000	1.000	66.00	59.80
8	1.000	1.000	1.000	97.00	67.80
9	0.000	0.000	0.000	81.00	59.20
10	0.000	0.000	0.000	75.00	60.40
11	0.000	0.000	0.000	76.00	59.10
12	0.000	0.000	0.000	83.00	60.60
13	–1.682	0.000	0.000	76.00	59.10
14	1.682	0.000	0.000	79.00	65.90
15	0.000	–1.682	0.000	85.00	60.00
16	0.000	1.682	0.000	97.00	60.70
17	0.000	0.000	–1.682	55.00	57.40
18	0.000	0.000	1.682	81.00	63.20
19	0.000	0.000	0.000	80.00	60.80
20	0.000	0.000	0.000	91.00	58.90

11-13. Un fabricante de herramientas de corte ha desarrollado dos ecuaciones empíricas para la vida de la herramienta en horas (y_1) y para el costo de la herramienta en dólares (y_2). Ambos modelos son funciones lineales de la dureza del acero (x_1) y de la fecha de fabricación (x_2). Las dos ecuaciones son

$$\hat{y}_1 = 10 + 5x_1 + 2x_2$$
$$\hat{y}_2 = 23 + 3x_1 + 4x_2$$

y ambas ecuaciones son válidas en el rango $-1.5 \leq x_i \leq 1.5$. El costo unitario de la herramienta debe estar abajo de \$27.50 y la vida debe exceder 12 horas para que el producto sea competitivo. ¿Existe algún conjunto de condiciones de operación factible para este proceso? ¿Dónde se recomendaría correr este proceso?

11-14. Se corre un diseño central compuesto en un proceso de deposición química por vapor y se obtienen los datos experimentales que se muestran a continuación. Se procesaron simultáneamente cuatro unidades experimentales en cada corrida del diseño, y las respuestas son la media y la varianza del espesor, calculadas en las cuatro unidades.

x_1	x_2	$\bar{y}$	s^2
–1	–1	360.6	6.689
1	–1	445.2	14.230
–1	1	412.1	7.088
1	1	601.7	8.586
1.414	0	518.0	13.130
–1.414	0	411.4	6.644
0	1.414	497.6	7.649
0	–1.414	397.6	11.740
0	0	530.6	7.836
0	0	495.4	9.306
0	0	510.2	7.956
0	0	487.3	9.127

a) Ajustar un modelo a la respuesta media. Analizar los residuales.
b) Ajustar un modelo a la respuesta varianza. Analizar los residuales.
c) Ajustar un modelo a $\ln(s^2)$. ¿Este modelo es superior al que se encontró en el inciso *b*?
d) Suponga que se quiere que el espesor medio esté en el intervalo 450 ± 25. Encontrar un conjunto de condiciones de operación que consiga este objetivo y que al mismo tiempo minimice la varianza.
e) Comentar los aspectos de la minimización de la varianza del inciso *d*. ¿Se ha minimizado también la varianza *total* del proceso?

11-15. Verificar que el diseño de primer orden ortogonal es también un diseño de primer orden rotable.

11-16. Demostrar que aumentar un diseño 2^k con n_c puntos centrales no afecta las estimaciones de β_i $(i = 1, 2, ..., k)$, pero que la estimación de la ordenada al origen β_0 es el promedio de las $2^k + n_c$ observaciones.

11-17. ***El diseño central compuesto rotable.*** Puede demostrarse que un diseño de segundo orden es rotable si $\Sigma_{u=1}^{n} x_{iu}^a x_{ju}^b = 0$ si *a* o *b* (o ambas) son impares y si $\Sigma_{u=1}^{n} x_{iu}^4 = 3\Sigma_{u=1}^{n} x_{iu}^2 x_{ju}^2$. Demostrar que para el diseño central compuesto estas condiciones llevan a $\alpha = (n_F)^{1/4}$ para la rotabilidad, donde n_F es el número de puntos en la porción factorial.

11-18. Verificar que el diseño central compuesto que se muestra abajo está separado en bloques ortogonales:

Bloque 1			Bloque 2			Bloque 3		
x_1	x_2	x_3	x_1	x_2	x_3	x_1	x_2	x_3
0	0	0	0	0	0	–1.633	0	0
0	0	0	0	0	0	1.633	0	0
1	1	1	1	1	–1	0	–1.633	0
1	–1	–1	1	–1	1	0	1.633	0
–1	–1	1	–1	1	1	0	0	–1.633
–1	1	–1	–1	–1	–1	0	0	1.633
						0	0	0
						0	0	0

11-19. ***Formación de bloques del diseño central compuesto.*** Considere un diseño central compuesto para $k = 4$ variables en dos bloques. ¿Puede encontrarse siempre un diseño rotable formado de bloques ortogonales?

11-20. ¿Cómo puede correrse un diseño hexagonal en dos bloques ortogonales?

11-21. En la tabla siguiente se muestra el rendimiento durante los cuatro primeros ciclos de un proceso químico. Las variables son el porcentaje de concentración (x_1) en los niveles 30, 31 y 32 y la temperatura (x_2) en 140, 142 y 144°F. Hacer el análisis utilizando métodos EVOP.

	Condiciones				
Ciclo	(1)	(2)	(3)	(4)	(5)
1	60.7	59.8	60.2	64.2	57.5
2	59.1	62.8	62.5	64.6	58.3
3	56.6	59.1	59.0	62.3	61.1
4	60.5	59.8	64.5	61.0	60.1

11-22. Suponga que se aproxima una superficie de respuesta con un modelo de orden d_1, tal como $\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \boldsymbol{\varepsilon}$, cuando la verdadera superficie está descrita por un modelo de orden $d_2 > d_1$; es decir, $E(\mathbf{y}) = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2$.

a) Demostrar que los coeficientes de regresión son sesgados, es decir, que $E(\hat{\boldsymbol{\beta}}_1) = \boldsymbol{\beta}_1 + \mathbf{A}\boldsymbol{\beta}_2$, donde $\mathbf{A} = (\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{X}_2$. Es común llamar a $\mathbf{A}$ la matriz alias.

b) Si $d_1 = 1$ y $d_2 = 2$, y se utiliza un diseño 2^k completo para ajustar el modelo, usar el resultado del inciso *a* para determinar la estructura de los alias.

c) Si $d_1 = 1$, $d_2 = 2$ y $k = 3$, encontrar la estructura de los alias, suponiendo que se usa un diseño 2^{3-1} para ajustar el modelo.

d) Si $d_1 = 1$, $d_2 = 2$ y $k = 3$, y se utiliza el diseño símplex del problema 11-3 para ajustar el modelo, determinar la estructura de los alias y comparar los resultados con el inciso *c*.

11-23. En un artículo ("Conozcamos todos el cuadrado latino", en *Quality Engineering*, vol. 1, pp. 453-465), J.S. Hunter ilustra algunos de los problemas asociados con los diseños factoriales fraccionados 3^{k-p}. El factor A es la cantidad de etanol agregada a un combustible estándar y el factor B representa la relación aire/combustible. La variable de respuesta es la emisión de monóxido de carbono (CO) en g/m³. El diseño se muestra abajo:

Diseño				Observaciones	
A	B	x_1	x_2	y	
0	0	−1	−1	66	62
1	0	0	−1	78	81
2	0	+1	−1	90	94
0	1	−1	0	72	67
1	1	0	0	80	81
2	1	+1	0	75	78
0	2	−1	+1	68	66
1	2	0	+1	66	69
2	2	+1	+1	60	58

Observe que se ha usado el sistema de notación de 0, 1 y 2 para representar los niveles bajo, intermedio y alto de los factores. Se ha usado también una "notación geométrica" de −1, 0 y +1. Se hacen dos réplicas de cada corrida del diseño.

a) Verificar que el modelo de segundo orden

$$\hat{y} = 78.5 + 4.5x_1 - 7.0x_2 - 4.5x_1^2 - 4.0x_2^2 - 9.0x_1x_2$$

es un modelo razonable para este experimento. Trazar los contornos de la concentración de CO en el espacio x_1, x_2.

b) Suponga ahora que en lugar de sólo dos factores, se usaron *cuatro* factores en un diseño factorial fraccionado 3^{4-2} y que se obtuvieron *exactamente* los mismos datos que en el inciso *a*. El diseño sería el siguiente:

Diseño								Observaciones	
A	B	C	D	x_1	x_2	x_3	x_4	y	
0	0	0	0	−1	−1	−1	−1	66	62
1	0	1	1	0	−1	0	0	78	81
2	0	2	2	+1	−1	+1	+1	90	94
0	1	2	1	−1	0	+1	0	72	67
1	1	0	2	0	0	−1	+1	80	81
2	1	1	0	+1	0	0	−1	75	78
0	2	1	2	−1	+1	0	+1	68	66
1	2	2	0	0	+1	+1	−1	66	69
2	2	0	1	+1	+1	−1	0	60	58

Confirmar que este diseño es un arreglo ortogonal L_9.

c) Calcular los promedios marginales de la respuesta CO en cada nivel de los cuatro factores *A, B, C* y *D*. Construir gráficas de estos promedios marginales e interpretar los resultados. ¿Los factores *C* y *D* parecen tener efectos grandes? ¿Estos factores tienen *en realidad* algún efecto sobre la emisión de CO? ¿Por qué su efecto aparente es grande?

d) El diseño del inciso *b* permite el ajuste del modelo

$$y=\beta_0+\sum_{i=1}^{4}\beta_i x_i+\sum_{i=1}^{4}\beta_{ii}x_i^2+\varepsilon$$

Suponga que el *verdadero* modelo es

$$y=\beta_0+\sum_{i=1}^{4}\beta_i x_i+\sum_{i=1}^{4}\beta_{ii}x_i^2+\sum\sum_{i<j}\beta_{ij}x_i x_j=\varepsilon$$

Demostrar que si las $\hat{\beta}_j$ representan las estimaciones de mínimos cuadrados de los coeficientes del modelo ajustado, entonces

$$E(\hat{\beta}_0)=\beta_0-\beta_{13}-\beta_{14}-\beta_{34}$$
$$E(\hat{\beta}_1)=\beta_1-(\beta_{23}+\beta_{24})/2$$
$$E(\hat{\beta}_2)=\beta_2-(\beta_{13}+\beta_{14}+\beta_{34})/2$$
$$E(\hat{\beta}_3)=\beta_3-(\beta_{12}+\beta_{24})/2$$
$$E(\hat{\beta}_4)=\beta_4-(\beta_{12}+\beta_{23})/2$$
$$E(\hat{\beta}_{11})=\beta_{11}-(\beta_{23}-\beta_{24})/2$$
$$E(\hat{\beta}_{22})=\beta_{22}+(\beta_{13}+\beta_{14}+\beta_{34})/2$$
$$E(\hat{\beta}_{33})=\beta_{33}-(\beta_{24}-\beta_{12})/2+\beta_{14}$$
$$E(\hat{\beta}_{44})=\beta_{44}-(\beta_{12}-\beta_{23})/2+\beta_{13}$$

¿Ayuda esto a explicar los efectos grandes de los factores *C* y *D* que se observaron gráficamente en el inciso *c*?

11-24. Suponga que es necesario diseñar un experimento para ajustar un modelo cuadrático en la región $-1 \leq x_i \leq +1$, $i = 1, 2$ sujeto a la restricción $x_1 + x_2 \leq 1$. Si se viola la restricción, el proceso no funcionará adecuadamente. No es posible hacer más de $n = 12$ corridas. Establecer los siguientes diseños:

a) Un modelo DCC "inscrito" con punto central en $x_1 = x_2 = 0$.
b) Un diseño factorial 3^2 "inscrito" con punto central en $x_1 = x_2 = -0.25$.
c) Un diseño optimal *D*.
d) Un diseño optimal *D* modificado que sea idéntico al del inciso *c*, pero con todas las réplicas de las corridas en el centro del diseño.
e) Evaluar el criterio $|(\mathbf{X}'\mathbf{X})^{-1}|$ para cada diseño.
f) Evaluar la eficiencia *D* para cada diseño en comparación con el diseño optimal *D* del inciso *c*.
g) ¿Qué diseño preferiría el lector? ¿Por qué?

11-25. Considere un diseño 2^3 para ajustar un modelo de primer orden.

a) Evaluar el criterio *D* $|(\mathbf{X}'\mathbf{X})^{-1}|$ para este diseño.
b) Evaluar el criterio *A* $\text{tr}(\mathbf{X}'\mathbf{X})^{-1}$ para este diseño.
c) Encontrar la varianza de predicción escalada máxima para este diseño. ¿Este diseño es optimal *G*?

11-26. Repetir el problema 11-25 utilizando un modelo de primer orden con las interacciones de dos factores.

11-27. Un ingeniero químico desea ajustar una curva de calibración para un nuevo procedimiento utilizado para medir la concentración de un ingrediente particular de un producto fabricado en sus instalaciones. Pueden prepararse 12 muestras, cuya concentración es conocida. El ingeniero quiere construir un modelo para las concentraciones medidas. Piensa que una curva de calibración lineal será adecuada para modelar la concentración medida como una función de las concentraciones conocidas; es decir, $y = \beta_0 + \beta_1 x + \varepsilon$, donde x es la concentración real. Están bajo consideración cuatro diseños experimentales. El diseño 1 consta de seis corridas con la concentración conocida 1 y seis corridas con la concentración conocida 10. El diseño 2 consta de cuatro corridas con las concentraciones 1, 5.5 y 10. El diseño 3 consta de tres corridas con las concentraciones 1, 4, 7 y 10. Por último, el diseño 4 consta de tres corridas con las concentraciones 1 y 10 y seis corridas con la concentración 5.5.

a) Graficar la varianza de predicción escalada para los cuatro diseños en la misma gráfica en el rango de la concentración $1 \leq x \leq 10$. ¿Qué diseño sería preferible?
b) Calcular el determinante de $(\mathbf{X}'\mathbf{X})^{-1}$ para cada diseño. ¿Qué diseño sería preferible de acuerdo con el criterio *D*?
c) Calcular la eficiencia *D* de cada diseño en comparación con el "mejor" diseño que se haya encontrado en el inciso *b*.
d) Para cada diseño, calcular la varianza de predicción promedio en el conjunto de puntos dado por $x = 1$, 1.5, 2, 2.5, ..., 10. ¿Qué diseño sería preferible de acuerdo con el criterio *V*?
e) Calcular la eficiencia *V* de cada diseño en comparación con el mejor diseño que se haya encontrado en el inciso *d*.
f) ¿Cuál es la eficiencia *G* de cada diseño?

11-28. Resolver de nuevo el problema 11-27, suponiendo que el modelo que el ingeniero quiere ajustar es cuadrático. Evidentemente, ahora sólo pueden considerarse los diseños 2, 3 y 4.

11-29. Un experimentador quiere correr un experimento de una mezcla de tres componentes. Las restricciones sobre las proporciones de los componentes son las siguientes:

$$0.2 \leq x_1 \leq 0.4$$
$$0.1 \leq x_2 \leq 0.3$$
$$0.4 \leq x_2 \leq 0.7$$

a) Establecer un experimento para ajustar un modelo cuadrático para mezclas. Usar $n = 14$ corridas, con cuatro réplicas. Usar el criterio *D*.
b) Trazar la región experimental.

c) Establecer un experimento para ajustar un modelo cuadrático para mezclas con $n = 12$ corridas, suponiendo que tres de estas corridas son réplicas. Usar el criterio *D*.

d) Comentar los dos diseños que se encontraron.

11-30. Myers y Montgomery [85a] describen un experimento con una mezcla de gasolina en el que intervienen tres componentes de la mezcla. No hay restricciones sobre las proporciones de la mezcla, y se usó el siguiente diseño con 10 corridas:

Punto del diseño	x_1	x_2	x_3	y, mi/gal
1	1	0	0	24.5, 25.1
2	0	1	0	24.8, 23.9
3	0	0	1	22.7, 23.6
4	$\frac{1}{2}$	$\frac{1}{2}$	0	25.1
5	$\frac{1}{2}$	0	$\frac{1}{2}$	24.3
6	0	$\frac{1}{2}$	$\frac{1}{2}$	23.5
7	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	24.8, 24.1
8	$\frac{2}{3}$	$\frac{1}{6}$	$\frac{1}{6}$	24.2
9	$\frac{1}{6}$	$\frac{2}{3}$	$\frac{1}{6}$	23.9
10	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{2}{3}$	23.7

a) ¿Qué tipo de diseño utilizaron los experimentadores?

b) Ajustar un modelo cuadrático para mezclas a los datos. ¿Es adecuado este modelo?

c) Graficar los contornos de la superficie de respuesta. ¿Qué mezcla se recomendaría para maximizar las millas por galón?

11-31. Considere el experimento del llenado de las botellas del ejemplo 6-1. Suponga que el porcentaje de carbonatación (*A*) es una variable de ruido (en unidades codificadas $\sigma_z^2 = 1$).

a) Ajustar el modelo de respuesta a estos datos. ¿Se trata de un problema de diseño robusto?

b) Encontrar el modelo de la media y el modelo de la varianza o bien la POE.

c) Encontrar un conjunto de condiciones que resulten en una desviación del llenado promedio tan próxima a cero como sea posible con varianza transmitida mínima.

11-32. Considere el experimento del problema 11-12. Suponga que la temperatura es una variable de ruido ($\sigma_z^2 = 1$ en unidades codificadas). Ajustar modelos de respuesta para las dos respuestas. ¿Se trata de un problema de diseño robusto con respecto a ambas respuestas? Encontrar un conjunto de condiciones que maximicen la conversión con la actividad entre 55 y 60 y que minimice la variabilidad transmitida por la temperatura.

11-33. Se ha corrido un experimento en un proceso que aplica un material de recubrimiento a una oblea. En cada corrida del experimento se fabricó una oblea y se midió varias veces el espesor del recubrimiento en varios sitios de la misma. Después se obtuvo la media y_1 y la desviación estándar y_2 de la medición del espesor. Los datos (adaptados de Box y Draper [16b]) se muestran en la tabla siguiente:

Corrida	Velocidad	Presión	Distancia	Media y_1	Desviación estándar y_2
1	−1.000	−1.000	−1.000	24.0	12.5
2	0.000	−1.000	−1.000	120.3	8.4
3	1.000	−1.000	−1.000	213.7	42.8
4	−1.000	0.000	−1.000	86.0	3.5
5	0.000	0.000	−1.000	136.6	80.4
6	1.000	0.000	−1.000	340.7	16.2
7	−1.000	1.000	−1.000	112.3	27.6

Corrida	Velocidad	Presión	Distancia	Media y_1	Desviación estándar y_2
8	0.000	1.000	–1.000	256.3	4.6
9	1.000	1.000	–1.000	271.7	23.6
10	–1.000	–1.000	0.000	81.0	0.0
11	0.000	–1.000	0.000	101.7	17.7
12	1.000	–1.000	0.000	357.0	32.9
13	–1.000	0.000	0.000	171.3	15.0
14	0.000	0.000	0.000	372.0	0.0
15	1.000	0.000	0.000	501.7	92.5
16	–1.000	1.000	0.000	264.0	63.5
17	0.000	1.000	0.000	427.0	88.6
18	1.000	1.000	0.000	730.7	21.1
19	–1.000	–1.000	1.000	220.7	133.8
20	0.000	–1.000	1.000	239.7	23.5
21	1.000	–1.000	1.000	422.0	18.5
22	–1.000	0.000	1.000	199.0	29.4
23	0.000	0.000	1.000	485.3	44.7
24	1.000	0.000	1.000	673.7	158.2
25	–1.000	1.000	1.000	176.7	55.5
26	0.000	1.000	1.000	501.0	138.9
27	1.000	1.000	1.000	1010.0	142.4

a) ¿Qué tipo de diseño utilizaron los experimentadores? ¿Es ésta una buena elección del diseño para ajustar un modelo cuadrático?

b) Construir los modelos para ambas respuestas.

c) Encontrar un conjunto de condiciones óptimas que resulten en una media tan grande como sea posible con la desviación estándar menor que 60.

11-34. ***Una variación del ejemplo 6-2.*** En el ejemplo 6-2 se encontró que una de las variables del proceso (B = presión) no era importante. Al eliminar esta variable se producen dos réplicas de un diseño 2^3. Los datos se muestran enseguida:

C	D	A (+)	A (–)	$\bar{y}$	s^2
–	–	45, 48	71, 65	57.75	121.19
+	–	68, 80	60, 65	68.25	72.25
–	+	43, 45	100, 104	73.00	1124.67
+	+	75, 70	86, 96	81.75	134.92

Suponer que C y D son factores controlables y que A es una variable de ruido.

a) Ajustar un modelo para la respuesta media.

b) Ajustar un modelo para la respuesta $ln(s^2)$.

c) Encontrar las condiciones de operación que resulten en la respuesta de rapidez de filtración media que exceda 75 con varianza mínima.

d) Comparar los resultados obtenidos con los del ejemplo 11-6, en el que se aplicó el enfoque de la transmisión del error. ¿Hasta qué punto son similares las dos respuestas?

12 Experimentos con factores aleatorios

A lo largo de gran parte de este libro se ha supuesto que los factores de un experimento son **factores fijos,** es decir, los niveles de los factores usados por el experimentador son los niveles de interés específico. La implicación de esto es, desde luego, que las inferencias estadísticas que se hacen acerca de estos factores se restringen a los niveles específicos estudiados. Es decir, si se investigan tres tipos de materiales, como en el experimento de la vida de la batería del ejemplo 5–1, las conclusiones sólo son válidas para esos tipos específicos de materiales. Una variante de esto ocurre cuando el factor o factores son **cuantitativos.** En estas situaciones, con frecuencia se usa un modelo de regresión que relaciona la respuesta con los factores para predecir la respuesta en la región que abarcan los niveles de los factores usados en el diseño experimental. Varios ejemplos de esto se presentaron en los capítulos 5 al 9. En general, cuando se trabaja con un efecto fijo, se dice que el **espacio inferencial** del experimento es el conjunto específico de los niveles de los factores investigados.

En algunas situaciones experimentales, los niveles de los factores se eligen al azar de una población más grande de niveles posibles, y el experimentador quiere sacar conclusiones acerca de la población completa de los niveles, no sólo de los que se usaron en el diseño experimental. En esta situación se dice que se trata de un **factor aleatorio.** Se empieza con una situación simple, un experimento con un solo factor en el que el factor es aleatorio y se usa esto para introducir el **modelo de efectos aleatorios** para el análisis de varianza y los **componentes de la varianza.** Los factores aleatorios ocurren también normalmente en experimentos factoriales, así como en otros tipos de experimentos. Este capítulo se enfoca en los métodos para el diseño y análisis de experimentos factoriales con factores aleatorios. En el capítulo 13 se presentarán los diseños anidados y de parcelas subdivididas, dos situaciones en las que es frecuente encontrar factores aleatorios en la práctica.

12-1 MODELO CON EFECTOS ALEATORIOS

Es común que un experimentador esté interesado en un factor que tiene un gran número de posibles niveles. Cuando el experimentador selecciona aleatoriamente a de estos niveles de la población de los niveles del factor, entonces se dice que el factor es **aleatorio.** Puesto que los niveles del factor utilizados realmente en el experimento se eligieron al azar, se hacen inferencias acerca de la población completa de los niveles del factor. Se supone que la población de los niveles del factor es de tamaño infinito o bien lo suficientemen-

te grande para considerarla infinita. No es frecuente encontrar situaciones en las que la población de los niveles del factor sea lo suficientemente pequeña para emplear el enfoque de una población finita. Referirse a Bennett y Franklin [9] y Searle y Fawcett [101] para una revisión del caso de una población finita.

El modelo estadístico lineal es

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij} \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, n \end{cases} \tag{12-1}$$

donde tanto τ_i como ε_{ij} son variables aleatorias. Si τ_i tiene varianza σ_τ^2 y es independiente de ε_{ij}, la varianza de cualquier observación es

$$V(y_{ij}) = \sigma_\tau^2 + \sigma^2$$

A las varianzas σ_τ^2 y σ^2 se les llama los **componentes de la varianza,** y al modelo (ecuación 12-1) se le llama **modelo de efectos aleatorios** o **de los componentes de la varianza.** Para probar hipótesis en este modelo se requiere que las $\{\varepsilon_{ij}\}$ sean NID$(0, \sigma^2)$, que las $\{\tau_i\}$ sean NID$(0, \sigma_\tau^2)$, y que τ_i y ε_{ij} sean independientes.[1]

La suma de cuadrados identidad

$$SS_T = SS_{\text{Tratamientos}} + SS_E \tag{12-2}$$

sigue siendo válida. Es decir, se hace la partición de la variabilidad total en las observaciones en un componente que mide la variación entre los tratamientos ($SS_{\text{Tratamientos}}$) y un componente que mide la variación dentro de los tratamientos (SS_E). Probar hipótesis acerca de los efectos de tratamientos individuales no tiene sentido, por lo que en su lugar se prueban hipótesis acerca del componente de la varianza σ_τ^2:

$$\begin{aligned} H_0&: \sigma_\tau^2 = 0 \\ H_1&: \sigma_\tau^2 > 0 \end{aligned} \tag{12-3}$$

Si $\sigma_\tau^2 = 0$, todos los tratamientos son idénticos; pero si $\sigma_\tau^2 > 0$, existe variabilidad entre los tratamientos. Como anteriormente, SS_E/σ^2 se distribuye como ji-cuadrada con $N - a$ grados de libertad y, bajo la hipótesis nula, $SS_{\text{Tratamientos}}/\sigma^2$ se distribuye como ji-cuadrada con $a - 1$ grados de libertad. Ambas variables aleatorias son independientes. Por lo tanto, bajo la hipótesis nula $\sigma_\tau^2 = 0$, el cociente

$$F_0 = \frac{\dfrac{SS_{\text{Tratamientos}}}{a-1}}{\dfrac{SS_E}{N-a}} = \frac{MS_{\text{Tratamientos}}}{MS_E} \tag{12-4}$$

se distribuye como F con $a - 1$ y $N - a$ grados de libertad. Sin embargo, es necesario examinar los cuadrados medios esperados para tener la descripción completa del procedimiento de prueba.

Considere

$$\begin{aligned} E(MS_{\text{Tratamientos}}) &= \frac{1}{a-1} E(SS_{\text{Tratamientos}}) = \frac{1}{a-1} E\left[\sum_{i=1}^{a} \frac{y_{i.}^2}{n} - \frac{y_{..}^2}{N}\right] \\ &= \frac{1}{a-1} E\left[\frac{1}{n}\sum_{i=1}^{a} \left(\sum_{j=1}^{n} \mu + \tau_i + \varepsilon_{ij}\right)^2 - \frac{1}{N}\left(\sum_{i=1}^{a}\sum_{j=1}^{n} \mu + \tau_i + \varepsilon_{ij}\right)^2\right] \end{aligned}$$

Cuando se eleva al cuadrado y se toma la función esperanza de las cantidades entre corchetes, se observa que los términos que incluyen a τ_i^2 son reemplazados por σ_τ^2 como $E(\tau_i) = 0$. Además, los términos que in-

[1] El supuesto de que las $\{\tau_i\}$ son variables aleatorias independientes implica que el supuesto usual de que $\sum_{i=1}^{a} \tau_i = 0$ del modelo de efectos fijos no se aplica al modelo de efectos aleatorios.

cluyen a $\varepsilon_{i.}^2$, $\varepsilon_{..}^2$ y $\Sigma_{i=1}^{a}\Sigma_{j=1}^{n}\tau_i^2$ son reemplazados por $n\sigma^2$, $an\sigma^2$ y $an^2\sigma_\tau^2$, respectivamente. Por otra parte, todos los productos cruzados que incluyen a τ_i y ε_{ij} tienen valor esperado cero. Esto lleva a

$$E(MS_{\text{Tratamientos}}) = \frac{1}{a-1}[N\mu^2 + N\sigma_\tau^2 + a\sigma^2 - N\mu^2 - n\sigma_\tau^2 - \sigma^2]$$

o

$$E(MS_{\text{Tratamientos}}) = \sigma^2 + n\sigma_\tau^2 \qquad (12\text{-}5)$$

De manera similar, puede demostrarse que

$$E(MS_E) = \sigma^2 \qquad (12\text{-}6)$$

Por los cuadrados medios esperados, se observa que bajo H_0 tanto el numerador como el denominador del estadístico de prueba (ecuación 12-4) son estimadores insesgados de σ^2, mientras que bajo H_1 el valor esperado del numerador es mayor que el valor esperado del denominador. Por lo tanto, H_0 deberá rechazarse para los valores de F_0 que sean muy grandes. Esto implica una región crítica de una cola superior, por lo que H_0 se rechaza si $F_0 > F_{\alpha, a-1, N-a}$.

El procedimiento de cálculo y el análisis de la tabla de varianza del modelo de efectos aleatorios son idénticos a los que se utilizaron en el caso de efectos fijos. Sin embargo, las conclusiones son muy diferentes, ya que se aplican a la población completa de los tratamientos.

Por lo general habrá interés en estimar los componentes de la varianza (σ^2 y σ_τ^2) del modelo. Al procedimiento que se usa para estimar σ^2 y σ_τ^2 se le llama **método del análisis de varianza,** ya que hace uso de las líneas de la tabla del análisis de varianza. El procedimiento consiste en igualar los cuadrados medios esperados con sus valores observados en la tabla del análisis de varianza y despejar los componentes de la varianza. Al igualar los cuadrados medios observados con los esperados en el modelo de efectos aleatorios con un solo factor, se obtiene

$$MS_{\text{Tratamientos}} = \sigma^2 + n\sigma_\tau^2$$

y

$$MS_E = \sigma^2$$

Por lo tanto, los estimadores de los componentes de la varianza son

$$\hat{\sigma}^2 = MS_E \qquad (12\text{-}7)$$

y

$$\hat{\sigma}_\tau^2 = \frac{MS_{\text{Tratamientos}} - MS_E}{n} \qquad (12\text{-}8)$$

Para tamaños de las muestras desiguales, se reemplaza n en la ecuación 12-8 con

$$n_0 = \frac{1}{a-1}\left[\sum_{i=1}^{a} n_i - \frac{\sum_{i=1}^{a} n_i^2}{\sum_{i=1}^{a} n_i}\right] \qquad (12\text{-}9)$$

En el método del análisis de varianza para estimar los componentes de la varianza no se requiere el supuesto de normalidad. Produce estimadores de σ^2 y σ_τ^2 que son los mejores estimadores cuadráticos insesgados (es decir, de todas las funciones cuadráticas insesgadas de las observaciones, estos estimadores tienen mínima varianza).

Ocasionalmente, el método del análisis de varianza produce una estimación negativa de uno de los componentes de la varianza. Evidentemente, los componentes de la varianza son por definición no negativos, por lo que la estimación negativa de un componente de la varianza se considera con cierta preocupación. Un curso de acción es aceptar la estimación y usarla como evidencia de que el verdadero valor del componente de la varianza es cero, suponiendo que la variación muestral llevó a la estimación negativa. Esto tiene un atractivo intuitivo, pero adolece de algunas dificultades teóricas. Por ejemplo, usar cero en lugar de la estimación negativa puede alterar las propiedades estadísticas de otras estimaciones. Otra alternativa es volver a estimar el componente de la varianza negativa utilizando un método que produzca siempre estimaciones no negativas. Una alternativa más es considerar la estimación negativa como evidencia de que el modelo lineal supuesto es incorrecto y examinar de nuevo el problema. El tratamiento completo de la estimación de los componentes de la varianza se ofrece en Searle [99a, b], Searle, Casella y McCullogh [100] y Burdick y Graybill [22].

EJEMPLO 12-1 ....

Una compañía textil fabrica un tejido en un gran número de telares. Le gustaría que los telares fueran homogéneos a fin de obtener un tejido de resistencia uniforme. El ingeniero del proceso sospecha que, además de la variación usual de la resistencia dentro de las muestras del tejido del mismo telar, puede haber también variaciones significativas en la resistencia entre un telar y otro. Para investigar esta posibilidad, el ingeniero selecciona cuatro telares al azar y hace cuatro determinaciones de la resistencia del tejido fabricado en cada telar. Este experimento se corre de manera aleatoria, y los datos obtenidos se muestran en la tabla 12-1. Se realiza el análisis de varianza, el cual se muestra en la tabla 12-2. Por el análisis de varianza se concluye que los telares de la planta difieren significativamente.

Los componentes de la varianza se estiman con $\hat{\sigma}^2 = 1.90$ y

$$\hat{\sigma}_\tau^2 = \frac{29.73 - 1.90}{4} = 6.96$$

Por lo tanto, la varianza de cualquier observación de la resistencia se estima con

$$\hat{\sigma}_\tau = \hat{\sigma}^2 + \hat{\sigma}_\tau^2 = 1.90 + 6.96 = 8.86.$$

La mayor parte de esta variabilidad es atribuible a las diferencias *entre* los telares.

....

Este ejemplo ilustra un uso importante de los componentes de la varianza: la separación de las diferentes fuentes de variabilidad que afectan un producto o sistema. El problema de la variabilidad de un producto se presenta con frecuencia en el control de calidad, y en muchas ocasiones es difícil aislar las

Tabla 12-1 Datos de la resistencia del ejemplo 12-1

	Observaciones				
Telares	1	2	3	4	$y_{i.}$
1	98	97	99	96	390
2	91	90	93	92	366
3	96	95	97	95	383
4	95	96	99	98	388
					$1527 = y_{..}$

Tabla 12-2 Análisis de varianza de los datos de la resistencia

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0	Valor P
Telares	89.19	3	29.73	15.68	<0.001
Error	22.75	12	1.90		
Total	111.94	15			

fuentes de la variabilidad. Por ejemplo, este estudio puede haber sido motivado por una gran variabilidad en la resistencia del tejido, como se ilustra en la figura 12-1*a*. En esta gráfica se presenta la salida del proceso (resistencia del tejido) modelado como una distribución normal con varianza $\hat{\sigma}_y^2 = 8.86$. (Ésta es la estimación de la varianza de cualquier observación de la resistencia del ejemplo 12-1.) Las especificaciones superior e inferior de la resistencia se muestran también en la figura 12-1*a*, y es relativamente inmediato ver que una proporción bastante grande de la salida del proceso se sale de las especificaciones (las áreas sombreadas de las colas de la figura 12-1*a*). El ingeniero del proceso se ha preguntado por qué es tan grande la cantidad de tejido defectuoso que debe desecharse, reelaborarse o degradarse a un producto de menor calidad. La respuesta es que la mayor parte de la variabilidad de la resistencia del producto es el resultado de las diferencias entre los telares. El desempeño irregular de los telares podría ser el resultado de una instalación incorrecta, un mantenimiento deficiente, una supervisión ineficaz, operadores sin la capacitación suficiente, fibra de entrada defectuosa, etcétera.

El ingeniero del proceso debe intentar ahora aislar las causas específicas de la diferencia en el desempeño de los telares. Si pudiera identificar y eliminar estas fuentes de variabilidad entre los telares, sería posible reducir considerablemente la varianza de la salida del proceso, quizá hasta $\hat{\sigma}_y^2 = 1.90$, la estimación del componente de la varianza dentro del telar (error) en el ejemplo 12-1. En la figura 12-1*b* se muestra la distribución normal de la resistencia de la fibra con $\hat{\sigma}_y^2 = 1.90$. Observe que la proporción del producto defectuoso en la salida se ha reducido radicalmente. Aun cuando es improbable que pueda eliminarse *toda* la variabilidad entre los telares, es claro que una reducción significativa en este componente de la varianza incrementaría sensiblemente la calidad de la fibra producida.

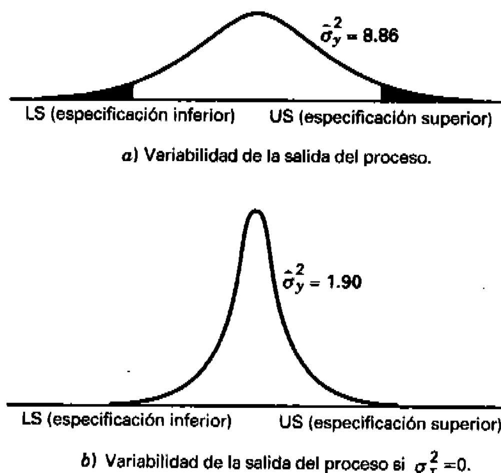


Figura 12-1 Salida del proceso en el problema de la resistencia de la fibra.

Es sencillo encontrar un intervalo de confianza para el componente de la varianza σ^2. Si las observaciones siguen una distribución normal e independiente, entonces $(N-a)MS_E/\sigma^2$ se distribuye como χ^2_{N-a}. Por lo tanto,

$$P\left[\chi^2_{1-(\alpha/2),N-a} \leq \frac{(N-a)MS_E}{\sigma^2} \leq \chi^2_{\alpha/2,N-a}\right] = 1-\alpha$$

y un intervalo de confianza de $100(1-\alpha)$ por ciento para σ^2 es

$$\frac{(N-a)MS_E}{\chi^2_{\alpha/2,N-a}} \leq \sigma^2 \leq \frac{(N-a)MS_E}{\chi^2_{1-(\alpha/2),N-a}} \quad (12\text{-}10)$$

Considere ahora el componente de la varianza σ^2_τ. El estimador puntual de σ^2_τ es

$$\hat{\sigma}^2_\tau = \frac{MS_{\text{Tratamientos}} - MS_E}{n}$$

La variable aleatoria $(a-1)MS_{\text{Tratamientos}}/(\sigma^2 + n\sigma^2_\tau)$ se distribuye como χ^2_{a-1}, y $(N-a)MS_E/\sigma^2$ se distribuye como χ^2_{N-a}. Por lo tanto, la distribución de probabilidad de $\hat{\sigma}^2_\tau$ es una combinación lineal de dos variables aleatorias ji-cuadrada, por ejemplo

$$u_1\chi^2_{a-1} - u_2\chi^2_{N-a}$$

donde

$$u_1 = \frac{\sigma^2 + n\sigma^2_\tau}{n(a-1)} \quad \text{y} \quad u_2 = \frac{\sigma^2}{n(N-a)}$$

Desafortunadamente, no puede obtenerse una expresión predeterminada para la distribución de esta combinación lineal de variables aleatorias ji-cuadrada. Por lo tanto, no es posible construir un intervalo de confianza exacto para σ^2_τ. En Graybill [50] y Searle [99a] se presentan procedimientos aproximados. Ver también la sección 12-7.

Es sencillo encontrar una expresión exacta para un intervalo de confianza del cociente $\sigma^2_\tau/(\sigma^2_\tau + \sigma^2)$. Se trata de un cociente con significado, ya que refleja la *proporción* de la varianza de una observación [recuerde que $V(y_{ij}) = \sigma^2_\tau + \sigma^2$] que es el resultado de las diferencias entre los tratamientos. Para desarrollar este intervalo de confianza en el caso de un diseño balanceado, observe que $MS_{\text{Tratamientos}}$ y MS_E son variables aleatorias independientes y, además, que puede demostrarse que

$$\frac{MS_{\text{Tratamientos}}/(n\sigma^2_\tau + \sigma^2)}{MS_E/\sigma^2} \sim F_{a-1,N-a}$$

Por lo tanto,

$$\left(F_{1-\alpha/2,a-1,N-a} \leq \frac{MS_{\text{Tratamientos}}}{MS_E}\frac{\sigma^2}{n\sigma^2_\tau + \sigma^2} \leq F_{\alpha/2,a-1,N-a}\right) = 1-\alpha \quad (12\text{-}11)$$

Al reordenar la ecuación 12-11, puede obtenerse la siguiente expresión:

$$P\left(L \leq \frac{\sigma^2_\tau}{\sigma^2} \leq U\right) = 1-\alpha \quad (12\text{-}12)$$

donde

$$L = \frac{1}{n}\left(\frac{MS_{\text{Tratamientos}}}{MS_E}\frac{1}{F_{\alpha/2,a-1,N-a}} - 1\right) \quad (12\text{-}13a)$$

y

$$U = \frac{1}{n}\left(\frac{MS_{\text{Tratamientos}}}{MS_E} \frac{1}{F_{1-\alpha/2,a-1,N-a}} - 1 \right) \tag{12-13b}$$

Observe que L y U son los límites de confianza inferior y superior del intervalo $100(1-\alpha)$ por ciento, respectivamente, del cociente σ_τ^2/σ^2. Por lo tanto, un intervalo de confianza de $100(1-\alpha)$ por ciento para $\sigma_\tau^2/(\sigma_\tau^2 + \sigma^2)$ es

$$\frac{L}{1+L} \leq \frac{\sigma_\tau^2}{\sigma_\tau^2 + \sigma^2} \leq \frac{U}{1+U} \tag{12-14}$$

Para ilustrar este procedimiento, se encontrará un intervalo de confianza de 95% de $\sigma_\tau^2/(\sigma_\tau^2 + \sigma^2)$ para los datos de la resistencia del ejemplo 12-1. Recuerde que $MS_{\text{Tratamientos}} = 29.73$, $MS_E = 1.90$, $a = 4$, $n = 4$, $F_{0.025,3,12} = 4.47$ y $F_{0.975,3,12} = 1/F_{0.025,12,3} = 1/14.34 = 0.070$. Por lo tanto, por las ecuaciones 12-13*a* y *b*,

$$L = \frac{1}{4}\left[\left(\frac{29.73}{1.90} \right)\left(\frac{1}{4.47} \right) - 1 \right] = 0.625$$

$$U = \frac{1}{4}\left[\left(\frac{29.73}{1.90} \right)\left(\frac{1}{0.070} \right) - 1 \right] = 54.883$$

y por la ecuación 12-12, el intervalo de confianza de 95% para $\sigma_\tau^2/(\sigma_\tau^2 + \sigma^2)$ es

$$\frac{0.625}{1.625} \leq \frac{\sigma_\tau^2}{\sigma_\tau^2 + \sigma^2} \leq \frac{54.883}{55.883}$$

o

$$0.39 \leq \frac{\sigma_\tau^2}{\sigma_\tau^2 + \sigma^2} \leq 0.98$$

Se concluye que la variabilidad entre los telares explica entre 39 y 98% de la varianza en la resistencia observada del tejido producido. Este intervalo de confianza es relativamente ancho debido al tamaño pequeño de la muestra que se usó en el experimento. Sin embargo, es evidente que la variabilidad entre los telares (σ_τ^2) no es insignificante.

12-2 DISEÑO FACTORIAL DE DOS FACTORES ALEATORIOS

Suponga que se tienen dos factores, A y B, y que ambos tienen un gran número de niveles de interés (como en la sección anterior, se supondrá que el número de niveles es infinito). Se escogerán al azar a niveles del factor A y b niveles del factor B, y estos niveles de los factores se incluirán en un diseño experimental factorial. Si el experimento se hace con n réplicas, las observaciones pueden representarse con el modelo lineal

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \varepsilon_{ijk} \qquad \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, b \\ k = 1, 2, \ldots, n \end{cases} \tag{12-15}$$

donde todos los parámetros del modelo, τ_i, β_j, $(\tau\beta)_{ij}$ y ε_{ijk}, son variables aleatorias independientes. También se supondrá que las variables aleatorias τ_i, β_j, $(\tau\beta)_{ij}$ y ε_{ijk} siguen una distribución normal con media cero y

varianzas dadas por $V(\tau_i) = \sigma_\tau^2$, $V(\beta_j) = \sigma_\beta^2$, $V[(\tau\beta)_{ij}] = \sigma_{\tau\beta}^2$ y $V(\varepsilon_{ijk}) = \sigma^2$. Por lo tanto, la varianza de cualquier observación es

$$V(y_{ijk}) = \sigma_\tau^2 + \sigma_\beta^2 + \sigma_{\tau\beta}^2 + \sigma^2 \tag{12-16}$$

y σ_τ^2, σ_β^2, $\sigma_{\tau\beta}^2$ y σ^2 son los **componentes de la varianza.** Las hipótesis que quieren probarse son $H_0\colon\sigma_\tau^2 = 0$, $H_0\colon\sigma_\beta^2 = 0$ y $H_0\colon\sigma_{\tau\beta}^2 = 0$. Observe la similitud con el modelo de efectos aleatorios de un solo factor.

Los cálculos numéricos del análisis de varianza se mantienen sin cambios; es decir, SS_A, SS_B, SS_{AB}, SS_T y SS_E se calculan como en el caso de efectos fijos. Sin embargo, para formar los estadísticos de prueba, deben examinarse los **cuadrados medios esperados.** Puede demostrarse que

$$\begin{aligned} E(MS_A) &= \sigma^2 + n\sigma_{\tau\beta}^2 + bn\sigma_\tau^2 \\ E(MS_B) &= \sigma^2 + n\sigma_{\tau\beta}^2 + an\sigma_\beta^2 \\ E(MS_{AB}) &= \sigma^2 + n\sigma_{\tau\beta}^2 \end{aligned} \tag{12-17}$$

y

$$E(MS_E) = \sigma^2$$

Por los cuadrados medios esperados se observa que el estadístico apropiado para probar la hipótesis de que no hay interacción, $H_0\colon\sigma_{\tau\beta}^2 = 0$, es

$$F_0 = \frac{MS_{AB}}{MS_E} \tag{12-18}$$

ya que bajo H_0 tanto el numerador como el denominador de F_0 tienen valor esperado σ^2, y sólo si H_0 es falsa $E(MS_{AB})$ es mayor que $E(MS_E)$. El cociente F_0 se distribuye como $F_{(a-1),ab(n-1)}$. De manera similar, para probar $H_0\colon\sigma_\tau^2 = 0$ se usaría

$$F_0 = \frac{MS_A}{MS_{AB}} \tag{12-19}$$

que se distribuye como $F_{a-1,(a-1)(b-1)}$, y para probar $H_0\colon\sigma_\beta^2 = 0$ el estadístico es

$$F_0 = \frac{MS_B}{MS_{AB}} \tag{12-20}$$

que se distribuye como $F_{b-1,(a-1)(b-1)}$. Todas estas pruebas son de una sola cola superior. Observe que estos estadísticos de prueba no son los mismos que se usarían si ambos factores A y B fuesen fijos. Los cuadrados medios esperados se usan siempre como guía para construir los estadísticos de prueba.

En muchos experimentos que incluyen factores aleatorios existe al menos interés tanto en estimar los componentes de la varianza como en la prueba de las hipótesis. Los componentes de la varianza pueden estimarse con el **método del análisis de varianza,** es decir, igualando los cuadrados medios observados de las líneas de la tabla del análisis de varianza con sus valores esperados y resolviendo para los componentes de la varianza. Se obtiene así

$$\begin{aligned} \hat\sigma^2 &= MS_E \\ \hat\sigma_{\tau\beta}^2 &= \frac{MS_{AB} - MS_E}{n} \\ \hat\sigma_\beta^2 &= \frac{MS_B - MS_{AB}}{an} \\ \hat\sigma_\tau^2 &= \frac{MS_A - MS_{AB}}{bn} \end{aligned} \tag{12-21}$$

como las estimaciones puntuales de los componentes de la varianza en el modelo de efectos aleatorios de dos factores. En la sección 12-7 se revisarán otros métodos para obtener estimaciones puntuales de los componentes de la varianza y los procedimientos para construir intervalos de confianza.

EJEMPLO 12-2 ··

Estudio de capacidad o aptitud de sistemas de medición

Con frecuencia se usan experimentos diseñados estadísticamente para investigar las fuentes de variabilidad que afectan a un sistema. Una aplicación industrial común es usar un experimento diseñado para estudiar los componentes de la variabilidad en un sistema de medición. Estos estudios se conocen comúnmente como **estudios de capacidad o aptitud de instrumentos de medición (calibradores)** o **estudios de repetibilidad y reproductibilidad (R&R) de instrumentos de medición (calibradores)**, ya que éstos son los componentes de la variabilidad de interés.

En la tabla 12-3 se muestra un experimento R&R de instrumentos de medición típico (de Montgomery [80a]). Se usa un instrumento o calibrador para medir una dimensión crítica de una pieza. Se han seleccionado 20 piezas del proceso de producción, y tres operadores escogidos al azar miden dos veces cada pieza con este calibrador. El orden en que se hacen las mediciones está completamente aleatorizado, por lo que se trata de un experimento factorial de dos factores en el que los factores del diseño son las piezas y los operadores, con dos réplicas. Las piezas y los operadores son factores aleatorios. Es válida la identidad del componente de la varianza de la ecuación 12-15; es decir,

$$\sigma_y^2 = \sigma_\tau^2 + \sigma_\beta^2 + \sigma_{\tau\beta}^2 + \sigma^2$$

donde σ_y^2 es la variabilidad total (que incluye la variabilidad debida a las diferentes piezas, la variabilidad debida a los diferentes operadores y la variabilidad debida al calibrador), σ_τ^2 es el componente de la varianza de las piezas, σ_β^2 es el componente de la varianza de los operadores, $\sigma_{\tau\beta}^2$ es el componente de la va-

Tabla 12-3 El experimento de la capacidad o aptitud del sistema de medición del ejemplo 12-2

Número de la pieza	Operador 1		Operador 2		Operador 3	
1	21	20	20	20	19	21
2	24	23	24	24	23	24
3	20	21	19	21	20	22
4	27	27	28	26	27	28
5	19	18	19	18	18	21
6	23	21	24	21	23	22
7	22	21	22	24	22	20
8	19	17	18	20	19	18
9	24	23	25	23	24	24
10	25	23	26	25	24	25
11	21	20	20	20	21	20
12	18	19	17	19	18	19
13	23	25	25	25	25	25
14	24	24	23	25	24	25
15	29	30	30	28	31	30
16	26	26	25	26	25	27
17	20	20	19	20	20	20
18	19	21	19	19	21	23
19	25	26	25	24	25	25
20	19	19	18	17	19	17

rianza que representa la interacción entre las piezas y los operadores, y σ^2 es el error experimental aleatorio. De manera típica, al componente de la varianza σ^2 se le llama la repetibilidad del instrumento de medición (calibrador), ya que puede considerarse que σ^2 refleja la variación obtenida cuando la misma pieza es medida por el mismo operador, y es común llamar a

$$\sigma_\beta^2 + \sigma_{\tau\beta}^2$$

la reproductibilidad del instrumento de medición (calibrador), ya que refleja la variabilidad adicional en el sistema de medición que resulta del uso del instrumento por parte del operador. Estos experimentos suelen realizarse con el objetivo de estimar los componentes de la varianza.

En la tabla 12-4 se muestra el análisis de varianza de este experimento. Los cálculos se realizaron utilizando la rutina Balanced ANOVA (análisis de varianza balanceado) de Minitab. Con base en los valores P, se concluye que el efecto de las piezas es grande, que los operadores quizá tengan un efecto pequeño y que no hay ninguna interacción significativa pieza-operador. La ecuación 12-21 puede usarse para estimar los componentes de la varianza de la siguiente manera:

$$\hat{\sigma}_\tau^2 = \frac{62.39 - 0.71}{(3)(2)} = 10.28$$

$$\hat{\sigma}_\beta^2 = \frac{1.31 - 0.71}{(20)(2)} = 0.015$$

$$\hat{\sigma}_{\tau\beta}^2 = \frac{0.71 - 0.99}{2} = -0.14$$

y

$$\hat{\sigma}^2 = 0.99$$

La parte inferior de la salida de Minitab de la tabla 12-4 contiene los cuadrados medios esperados del modelo aleatorio, con los números entre paréntesis representando los componentes de la varianza [(4) representa σ^2, (3) representa $\sigma_{\tau\beta}^2$, etc.]. Se presentan también las estimaciones de los componentes de la varianza, junto con el término del error que se utilizó para probar ese componente de la varianza en el análisis de varianza. Más adelante se estudiará la terminología **modelo no restringido;** ésta no tiene relevancia en los modelos aleatorios.

Observe que la estimación de uno de los componentes de la varianza, $\hat{\sigma}_{\tau\beta}^2$, es negativa. Desde luego, esto no tiene sentido, ya que por definición las varianzas son no negativas. Desafortunadamente, pueden obtenerse estimaciones negativas de los componentes de la varianza cuando se usa el método de estimación del análisis de varianza (lo cual se considera una de sus desventajas). Existen varias maneras de abordar esta situación. Una posibilidad es suponer que la estimación negativa significa que el componente de la varianza en realidad es cero y simplemente se hace cero, dejando sin cambios las demás estimaciones no negativas. Otro enfoque es estimar los componentes de la varianza con un método que asegure estimaciones no negativas (este enfoque se revisará brevemente en la sección 12-7). Por último, podría observarse que el valor P del término de interacción de la tabla 12-4 es muy grande, tomar esto como evidencia de que $\sigma_{\tau\beta}^2$ es en realidad cero (es decir, que no hay efecto de interacción) y ajustar un **modelo reducido** de la forma

$$y_{ijk} = \mu + \tau_i + \beta_j + \varepsilon_{ijk}$$

que no incluye el término de interacción. Éste es un enfoque relativamente sencillo y que con frecuencia funciona casi tan bien como los métodos más elaborados.

Tabla 12-4 Análisis de varianza balanceado (Balanced ANOVA de Minitab) del ejemplo 12-2

Análisis de varianza (diseños balanceados)

```
Factor     Type Levels Values
part     random     20      1      2      3      4      5      6      7
                            8      9     10     11     12     13     14
                           15     16     17     18     19     20
operator random      3      1      2      3

Analysis of Variance for y

Source            DF         SS        MS       F       P
part              19   1185.425    62.391   87.65   0.000
operator           2      2.617     1.308    1.84   0.173
part*operator     38     27.050     0.712    0.72   0.861
Error             60     59.500     0.992
Total            119   1274.592

Source               Variance  Error   Expected Mean Square for Each Term
                    component  term    (using unrestricted model)
 1 part               10.2798    3      (4) + 2(3) + 6(1)
 2 operator            0.0149    3      (4) + 2(3) + 40(2)
 3 part*operator      -0.1399    4      (4) + 2(3)
 4 Error               0.9917           (4)
```

Tabla 12-5 Análisis de varianza del modelo reducido, ejemplo 12-2

Análisis de varianza (diseños balanceados)

```
Factor     Type Levels Values
part     random     20    1    2    3    4    5    6    7
                          8    9   10   11   12   13   14
                         15   16   17   18   19   20
operator random      3    1    2    3

Analysis of Variance for y

Source      DF        SS        MS       F       P
part        19  1185.425    62.391   70.64   0.000
operator     2     2.617     1.308    1.48   0.232
Error       98    86.550     0.883
Total      119  1274.592

Source          Variance   Error   Expected Mean Square for Each Term
                component  term    (using unrestricted model)
 1 part          10.2513    3      (3) + 6(1)
 2 operator       0.0106    3      (3) + 40(2)
 3 Error          0.8832           (3)
```

En la tabla 12-5 se muestra el análisis de varianza del modelo reducido. Puesto que no hay término de interacción, los dos efectos principales se prueban contra el término del error, y las estimaciones de los componentes de la varianza son

$$\hat{\sigma}_\tau^2 = \frac{62.39 - 0.88}{(3)(2)} = 10.25$$

$$\hat{\sigma}_\beta^2 = \frac{1.31 - 0.88}{(20)(2)} = 0.0108$$

$$\hat{\sigma}^2 = 0.88$$

Por último, la varianza del calibrador podría estimarse como la suma de las estimaciones de los componentes de la varianza $\hat{\sigma}^2$ y $\hat{\sigma}_\beta^2$ como

$$\begin{aligned}\hat{\sigma}^2_{\text{calibrador}} &= \hat{\sigma}^2 + \hat{\sigma}_\beta^2 \\ &= 0.88 + 0.0108 \\ &= 0.8908\end{aligned}$$

La variabilidad del calibrador parece ser pequeña en comparación con la variabilidad del producto. Se trata generalmente de una situación deseable, la cual implica que el calibrador tiene la capacidad de distinguir entre las diferentes gradaciones del producto.

12-3 MODELO MIXTO CON DOS FACTORES

Se considera ahora la situación en que uno de los factores, A, está fijo y el otro, B, es aleatorio. Se le llama análisis de varianza del **modelo mixto.** El modelo estadístico lineal es

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \varepsilon_{ijk} \qquad \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, b \\ k = 1, 2, \ldots, n \end{cases} \tag{12-22}$$

Aquí τ_i es un efecto fijo, β_j es un efecto aleatorio, se supone que la interacción $(\tau\beta)_{ij}$ es un efecto aleatorio y ε_{ijk} es un error aleatorio. Se supone también que las $\{\tau_i\}$ son efectos fijos tales que $\Sigma_{i=1}^{a}\tau_i = 0$ y que β_j es una variable aleatoria NID$(0, \sigma_\beta^2)$. El efecto de la interacción, $(\tau\beta)_{ij}$, es una variable aleatoria normal con media 0 y varianza $[(a-1)/a]\sigma_{\tau\beta}^2$; sin embargo, la operación suma del componente de la interacción en el rango del factor fijo es igual a cero. Es decir,

$$\sum_{i=1}^{a} (\tau\beta)_{ij} = (\tau\beta)_{.j} = 0 \qquad j = 1, 2, \ldots, b$$

Esta restricción implica que algunos elementos de la interacción en diferentes niveles del factor fijo no son independientes. De hecho, puede demostrarse (ver el problema 12-25) que

$$\text{Cov}[(\tau\beta)_{ij}, (\tau\beta)_{i'j}] = -\frac{1}{a}\sigma_{\tau\beta}^2 \qquad i \neq i'$$

La covarianza entre $(\tau\beta)_{ij}$ y $(\tau\beta)_{ij'}$ para $j \neq j'$ es cero, y el error aleatorio ε_{ijk} es NID$(0, \sigma^2)$. Puesto que la suma de los efectos de la interacción en los niveles del factor fijo es igual a cero, a esta versión del modelo mixto con frecuencia se le llama **modelo restringido.**

En este modelo la varianza de $(\tau\beta)_{ij}$ se define como $[(a-1)/a]\sigma_{\tau\beta}^2$ en vez de como $\sigma_{\tau\beta}^2$ para simplificar los cuadrados medios esperados. El supuesto $(\tau\beta)_{.j} = 0$ también tiene un efecto sobre los cuadrados medios esperados, los cuales puede demostrarse que son

$$\begin{aligned} E(MS_A) &= \sigma^2 + n\sigma_{\tau\beta}^2 + \frac{bn\sum_{i=1}^{a}\tau_i^2}{a-1} \\ E(MS_B) &= \sigma^2 + an\sigma_\beta^2 \\ E(MS_{AB}) &= \sigma^2 + n\sigma_{\tau\beta}^2 \end{aligned} \qquad (12\text{-}23)$$

y

$$E(MS_E) = \sigma^2$$

Por lo tanto, el estadístico de prueba apropiado para probar que las medias de los efectos del factor fijo son iguales, o $H_0{:}\tau_i = 0$, es

$$F_0 = \frac{MS_A}{MS_{AB}}$$

que tiene la distribución de referencia $F_{a-1,(a-1)(b-1)}$. Para probar $H_0 : \sigma_\beta^2$, el estadístico de prueba es

$$F_0 = \frac{MS_B}{MS_E}$$

con la distribución de referencia $F_{b-1,\, ab(n-1)}$. Por último, para probar la hipótesis de la interacción $H_0 : \sigma_{\tau\beta}^2 = 0$, se usaría

$$F_0 = \frac{MS_{AB}}{MS_E}$$

que tiene la distribución de referencia $F_{(a-1)(b-1),\, ab(n-1)}$.

En el modelo mixto es posible estimar los efectos del factor fijo como

$$\begin{aligned} \hat{\mu} &= \bar{y}_{...} \\ \hat{\tau}_i &= \bar{y}_{i..} - \bar{y}_{...} \qquad i = 1, 2, \ldots, a \end{aligned} \qquad (12\text{-}24)$$

Los componentes de la varianza $\sigma_\beta^2, \sigma_{\tau\beta}^2$ y σ^2 pueden estimarse aplicando el método del análisis de varianza. Al eliminar la primera ecuación de las ecuaciones 12-23 quedan tres ecuaciones con tres incógnitas, cuyas soluciones son

$$\hat{\sigma}_\beta^2 = \frac{MS_B - MS_E}{an} \qquad \hat{\sigma}_{\tau\beta}^2 = \frac{MS_{AB} - MS_E}{n} \tag{12-25}$$

y

$$\hat{\sigma}^2 = MS_E$$

Este enfoque general puede emplearse para estimar los componentes de la varianza en *cualquier* modelo mixto. Después de eliminar los cuadrados medios que contienen factores fijos, siempre quedará un sistema de ecuaciones que puede resolverse para los componentes de la varianza.

En los modelos mixtos, el experimentador puede tener interés en probar hipótesis o en construir intervalos de confianza para las medias de tratamientos individuales del factor fijo. Al utilizar estos procedimientos, deberá tenerse cuidado de usar el error estándar apropiado de la media de los tratamientos. El error estándar de la media de los tratamientos del efecto fijo es

$$\left[\frac{\text{Cuadrado medio para probar el efecto fijo}}{\text{Número de observaciones en la media de cada tratamiento}}\right]^{1/2} = \sqrt{\frac{MS_{AB}}{bn}}$$

Observe que esto es simplemente el error estándar que se usaría si éste fuera un modelo con efectos fijos, salvo porque MS_E se ha reemplazado con el cuadrado medio que se usó en la prueba de la hipótesis.

EJEMPLO 12-3

Retomando el experimento de la capacidad o aptitud del sistema de medición

Considere de nuevo el experimento R&R del calibrador descrito en el ejemplo 12-2. Suponga ahora que sólo tres operadores usan este calibrador, de tal modo que los operadores son un factor fijo. Sin embargo, puesto que las piezas se eligen al azar, se trata ahora de un experimento con un modelo mixto.

El análisis de varianza del modelo mixto se muestra en la tabla 12-6. Los cálculos se realizaron utilizando la rutina Balanced ANOVA (análisis de varianza balanceado) de Minitab. Se especificó el uso del modelo restringido en el análisis de Minitab, el cual generó también los cuadrados medios esperados para este modelo. En la salida de Minitab, la cantidad $Q[2]$ indica una expresión cuadrática que incluye al operador del efecto de factor fijo. Es decir, $Q[2] = \Sigma_{j=1}^{b}\beta_j^2/(b-1)$. Las conclusiones son similares al ejemplo 12-2. Los componentes de la varianza pueden estimarse con la ecuación 12-25 como

$$\hat{\sigma}_{\text{Piezas}}^2 = \frac{MS_{\text{Piezas}} - MS_E}{an} = \frac{62.39 - 0.99}{(3)(2)} = 10.23$$

$$\hat{\sigma}_{\text{Piezas} \times \text{operadores}}^2 = \frac{MS_{\text{Piezas} \times \text{operadores}} - MS_E}{n} = \frac{0.71 - 0.99}{2} = -0.14$$

$$\hat{\sigma}^2 = MS_E = 0.99$$

Estos resultados también se muestran en la salida de Minitab. De nueva cuenta, resulta una estimación negativa del componente de la varianza de la interacción. Un curso de acción apropiado sería ajustar un

Tabla 12-6 Análisis de varianza (Minitab) del modelo mixto del ejemplo 12-3. Se supone el modelo restringido

Análisis de varianza (diseños balanceados)

```
Factor     Type Levels Values
part     random     20     1     2     3     4     5     6     7
                           8     9    10    11    12    13    14
                          15    16    17    18    19    20
operator fixed       3     1     2     3

Analysis of Variance for y

Source            DF        SS        MS        F        P
part              19  1185.425    62.391    62.92    0.000
operator           2     2.617     1.308     1.84    0.173
part*operator     38    27.050     0.712     0.72    0.861
Error             60    59.500     0.992
Total            119  1274.592

Source                Variance  Error  Expected Mean Square for Each Term
                     component  term   (using restricted model)
 1 part                10.2332     4    (4) + 6(1)
 2 operator                        3    (4) + 2(3) + 40Q[2]
 3 part*operator       -0.1399     4    (4) + 2(3)
 4 Error                0.9917          (4)
```

modelo reducido, como se hizo en el ejemplo 12-2. En el caso de un modelo mixto con dos factores, esto lleva a los mismos resultados del ejemplo 12-2.

Modelos mixtos alternativos

Se han propuesto varias versiones diferentes del modelo mixto. Estos modelos difieren de la versión restringida del modelo mixto estudiado anteriormente en los supuestos establecidos acerca de los componentes aleatorios. A continuación se revisa brevemente uno de estos modelos alternativos.

Considere el modelo

$$y_{ijk} = \mu + \alpha_i + \gamma_j + (\alpha\gamma)_{ij} + \varepsilon_{ijk}$$

donde α_i ($i = 1, 2, \ldots, a$) son efectos fijos tales que $\Sigma_{i=1}^{a}\alpha_i = 0$ y γ_j, $(\alpha\gamma)_{ij}$ y ε_{ijk} son variables aleatorias no correlacionadas que tienen media cero y varianzas $V(\gamma_j) = \sigma_\gamma^2$, $V[(\alpha\gamma)_{ij}] = \sigma_{\alpha\gamma}^2$ y $V(\varepsilon_{ijk}) = \sigma^2$. Observe que aquí no se usa la restricción impuesta anteriormente sobre el efecto de la interacción; por consiguiente, a esta versión del modelo mixto se le llama con frecuencia **modelo mixto no restringido.**

Es posible demostrar que los cuadrados medios esperados para este modelo son (referirse al material suplementario del texto de este capítulo)

$$
\begin{aligned}
E(MS_A) &= \sigma^2 + n\sigma_{\alpha\gamma}^2 + \frac{bn\sum_{i=1}^{a}\alpha_i^2}{a-1} \\
E(MS_B) &= \sigma^2 + n\sigma_{\alpha\gamma}^2 + an\sigma_\gamma^2 \\
E(MS_{AB}) &= \sigma^2 + n\sigma_{\alpha\gamma}^2
\end{aligned}
\qquad (12\text{-}26)
$$

y

$$E(MS_E) = \sigma^2$$

Al comparar estos cuadrados medios esperados con los de la ecuación 12-23, se observa que la única diferencia evidente es la presencia del componente de la varianza $\sigma_{\alpha\gamma}^2$ en el cuadrado medio esperado del efecto aleatorio. (En realidad, hay otras diferencias debido a las definiciones diferentes de la varianza del efecto de la interacción en los dos modelos.) Por consiguiente, se probaría la hipótesis de que el componente de la varianza del efecto aleatorio es igual a cero (H_0:$\sigma_{\alpha\gamma}^2 = 0$) usando el estadístico

$$F_0 = \frac{MS_B}{MS_{AB}}$$

en contraste con probar $H_0: \sigma_\beta^2$ con $F_0 = MS_B/MS_E$ en el modelo restringido. La prueba deberá ser más conservadora cuando se emplee este modelo porque por lo general MS_{AB} será mayor que MS_E.

Los parámetros de los dos modelos guardan una relación cercana. De hecho, puede demostrarse que

$$
\begin{aligned}
\beta_j &= \gamma_j + (\overline{\alpha\gamma})_{.j} \\
(\tau\beta)_{ij} &= (\alpha\gamma)_{ij} + (\overline{\alpha\gamma})_{.j} \\
\sigma_\gamma^2 &= \sigma_\beta^2 + \frac{1}{a}\sigma_{\alpha\gamma}^2
\end{aligned}
$$

y

$$\sigma_{\tau\beta}^2 = \sigma_{\alpha\gamma}^2$$

Puede usarse el método del análisis de varianza para estimar los componentes de la varianza. Con referencia a los cuadrados medios esperados, se encuentra que el único cambio de las ecuaciones 12-25 es que

$$\hat{\sigma}_\gamma^2 = \frac{MS_B - MS_{AB}}{an} \tag{12-27}$$

Estos dos modelos son casos especiales del modelo mixto propuesto por Scheffé [98b, d]. En este modelo se supone que las observaciones pueden representarse con

$$y_{ijk} = m_{ij} + \varepsilon_{ijk} \qquad \begin{cases} i = 1, 2, \dots, a \\ j = 1, 2, \dots, b \\ k = 1, 2, \dots, n \end{cases}$$

donde m_{ij} y ε_{ijk} son variables aleatorias independientes. La estructura de m_{ij} es

$$m_{ij} = \mu + \tau_i + b_j + c_{ij}$$

$$E(m_{ij}) = \mu + \tau_i$$

$$\sum_{i=1}^{a} \tau_i = 0$$

y

$$c_{.j} = 0 \qquad j = 1, 2, \dots, b$$

Las varianzas y covarianzas de b_j y c_{ij} se expresan a través de las covarianzas de las m_{ij}. Además, los parámetros de los efectos aleatorios en otras formulaciones del modelo mixto pueden relacionarse con b_j y c_{ij}. El análisis estadístico del modelo de Scheffé es idéntico al del modelo restringido tratado aquí, salvo porque, en general, el estadístico MS_A/MS_{AB} no siempre se distribuye como F cuando H_0: $\tau_i = 0$ es verdadera.

A la luz de esta multiplicidad de modelos mixtos, una pregunta lógica es: ¿qué modelo deberá usarse? La mayoría de los especialistas en estadística prefieren el modelo restringido, mismo que se encuentra con mayor frecuencia en la literatura del tema. El modelo restringido es en realidad un poco más general que el no restringido, ya que en el primero la covarianza entre dos observaciones del mismo nivel del factor aleatorio puede ser positiva o negativa, mientras que en el segundo esta covarianza sólo puede ser positiva. Si la estructura correlativa de los componentes aleatorios no es grande, entonces cualquiera de los dos modelos mixtos es apropiado, y sólo hay diferencias menores entre ellos. Cuando se haga referencia más adelante a los modelos mixtos, se supondrá la estructura del modelo restringido. Sin embargo, si hay correlaciones grandes en los datos, entonces quizá deba emplearse el modelo de Scheffé. La elección del modelo deberá ser siempre dictada por los datos. El artículo de Hocking [56] es un resumen claro de diferentes modelos mixtos.

EJEMPLO 12-4 ....

El modelo no restringido

Algunos paquetes de software de computadora tienen soporte para un solo modelo mixto. Minitab soporta tanto el modelo restringido como el no restringido, aun cuando la selección por omisión es el modelo no restringido. En la tabla 12-7 se muestra la salida de Minitab para el experimento del ejemplo 12-3 utilizando el modelo no restringido. Observe que los cuadrados medios esperados concuerdan con los de la ecuación 12-26. Las conclusiones son idénticas a las del análisis del modelo restringido, y las estimaciones de los componentes de la varianza son muy similares.

....

Tabla 12-7 Análisis del experimento del ejemplo 12-3 utilizando el modelo restringido

Análisis de varianza (diseños balanceados)

Factor	Type	Levels	Values						
part	random	20	1	2	3	4	5	6	7
			8	9	10	11	12	13	14
			15	16	17	18	19	20	
operator	fixed	3	1	2	3				

Analysis of Variance for y

Source	DF	SS	MS	F	P
part	19	1185.425	62.391	87.65	0.000
operator	2	2.617	1.308	1.84	0.173
part*operator	38	27.050	0.712	0.72	0.861
Error	60	59.500	0.992		
Total	119	1274.592			

Source	Variance component	Error term	Expected Mean Square for Each Term (using unrestricted model)
1 part	10.2798	3	(4) + 2(3) + 6(1)
2 operator		3	(4) + 2(3) + Q[2]
3 part*operator	-0.1399	4	(4) + 2(3)
4 Error	0.9917		(4)

12-4 DETERMINACIÓN DEL TAMAÑO DE LA MUESTRA CON EFECTOS ALEATORIOS

Pueden usarse las curvas de operación característica del apéndice para determinar el tamaño de la muestra en experimentos con factores aleatorios. Se empieza con el modelo de efectos aleatorios con un solo factor de la sección 12-1. La probabilidad del error tipo II para el modelo de efectos aleatorios es

$$\begin{aligned}\beta &= 1 - P\{\text{Rechazar } H_0 \mid H_0 \text{ es falsa}\}\\ &= 1 - P\{F_0 > F_{\alpha,a-1,N-a} \mid \sigma_\tau^2 > 0\}\end{aligned} \tag{12-28}$$

De nueva cuenta se requiere la distribución del estadístico de prueba $F_0 = MS_{\text{Tratamientos}}/MS_E$ bajo la hipótesis alternativa. Puede demostrarse que si H_1 es verdadera ($\sigma_\tau^2 > 0$), la distribución de F_0 es F central con $a - 1$ y $N - a$ grados de libertad.

Puesto que la probabilidad del error tipo II del modelo de efectos aleatorios se basa en la distribución F central usual, podrían usarse las tablas de la distribución F del apéndice para evaluar la ecuación 12-28. Sin embargo, es más simple determinar la sensibilidad de la prueba mediante el uso de las curvas de operación característica. En la parte IV del apéndice se presenta una serie de estas curvas para varios valores de los grados de libertad del numerador, de los grados de libertad del denominador y α de 0.05 y 0.01. En estas curvas se grafica la probabilidad del error tipo II contra el parámetro λ, donde

$$\lambda = \sqrt{1 + \frac{n\sigma_\tau^2}{\sigma^2}} \tag{12-29}$$

Observe que λ incluye dos parámetros desconocidos, σ^2 y σ_τ^2. Quizá pueda estimarse σ_τ^2 si se tiene una idea acerca de cuánta variabilidad de la población de tratamientos es importante detectar. Puede escogerse una estimación de σ^2 recurriendo a la experiencia previa o discrecionalmente. En ocasiones es útil definir el valor de σ_τ^2 que quiere detectarse en términos del cociente σ_τ^2/σ^2.

EJEMPLO 12-5 ··

Suponga que se han seleccionado cinco tratamientos al azar con seis observaciones por tratamiento y que $\alpha = 0.05$, y quiere determinarse la potencia de la prueba cuando σ_τ^2 es igual a σ^2. Puesto que $a = 5, n = 6$ y $\sigma_\tau^2 = \sigma^2$, puede calcularse

$$\lambda = \sqrt{1 + 6(1)} = 2.646$$

Por la curva de operación característica con $a - 1 = 4, N - a = 25$ grados de libertad y $\alpha = 0.05$, se encuentra que

$$\beta \simeq 0.20$$

y por lo tanto la potencia es de aproximadamente 0.80.

··

También puede usarse el incremento porcentual en la desviación estándar de un método de observación para determinar el tamaño de la muestra. Si los tratamientos son homogéneos, entonces la desviación estándar de una observación seleccionada al azar es σ. Sin embargo, si los tratamientos son diferentes, la desviación estándar de una observación elegida al azar es

$$\sqrt{\sigma^2 + \sigma_\tau^2}$$

Si P es el incremento porcentual fijo en la desviación estándar de una observación más allá del cual se desea rechazar la hipótesis nula, entonces

$$\frac{\sqrt{\sigma^2+\sigma_\tau^2}}{\sigma}=1+0.01P$$

o

$$\frac{\sigma_\tau^2}{\sigma^2}=(1+0.01P)^2-1$$

Por lo tanto, utilizando la ecuación 12-29, se encuentra que

$$\lambda=\sqrt{1+\frac{n\sigma_\tau^2}{\sigma^2}}=\sqrt{1+n[(1+0.01P)^2-1]} \tag{12-30}$$

Para una P dada pueden usarse las curvas de operación característica de la parte VI del apéndice para encontrar el tamaño de la muestra deseado.

También pueden usarse las curvas de operación característica para determinar el tamaño de la muestra del modelo de efectos aleatorios con dos factores y del modelo mixto. Se utiliza la parte VI del apéndice para el modelo de efectos aleatorios. El parámetro λ, los grados de libertad del numerador y los grados de libertad del denominador se muestran en la mitad superior de la tabla 12-8. Para el modelo mixto deben usarse las partes V y VI del apéndice. Los valores apropiados de Φ^2 y λ se muestran en la mitad inferior de la tabla 12-8.

Tabla 12-8 Parámetros de las curvas de operación característica de las tablas V y VI del apéndice para los modelos con dos factores de efectos aleatorios y mixto

El modelo de efectos aleatorios			
Factor	λ	Grados de libertad del numerador	Grados de libertad del denominador
A	$\sqrt{1+\frac{bn\sigma_\tau^2}{\sigma^2+n\sigma_{\tau\beta}^2}}$	$a-1$	$(a-1)(b-1)$
B	$\sqrt{1+\frac{an\sigma_\beta^2}{\sigma^2+n\sigma_{\tau\beta}^2}}$	$b-1$	$(a-1)(b-1)$
AB	$\sqrt{1+\frac{n\sigma_{\tau\beta}^2}{\sigma^2}}$	$(a-1)(b-1)$	$ab(n-1)$

El modelo mixto				
Factor	Parámetro	Grados de libertad del numerador	Grados de libertad del denominador	Parte del apéndice
A (fijo)	$\Phi^2=\frac{bn\sum_{i=1}^{a}\tau_i^2}{a[\sigma^2+n\sigma_{\tau\beta}^2]}$	$a-1$	$(a-1)(b-1)$	V
B (aleatorio)	$\lambda=\sqrt{1+\frac{an\sigma_\beta^2}{\sigma^2}}$	$b-1$	$ab(n-1)$	VI
AB	$\lambda=\sqrt{1+\frac{n\sigma_{\tau\beta}^2}{\sigma^2}}$	$(a-1)(b-1)$	$ab(n-1)$	VI

12-5 REGLAS PARA LOS CUADRADOS MEDIOS ESPERADOS

Una parte importante de cualquier problema de diseño experimental es la realización del análisis de varianza. Esto implica determinar la suma de cuadrados de cada componente del modelo y el número de grados de libertad asociados con cada suma de cuadrados. Después, para construir los estadísticos de prueba apropiados, deben determinarse los cuadrados medios esperados. En situaciones de diseño complejas, particularmente las que incluyen modelos aleatorios o mixtos, con frecuencia es útil contar con un procedimiento formal para este proceso.

Se presentará un conjunto de reglas para anotar los cuadrados medios esperados en cualquier experimento factorial balanceado, anidado[2] o factorial anidado. (Observe que los arreglos parcialmente balanceados, como los cuadrados latinos y los diseños de bloques incompletos, se excluyen explícitamente.) Estas reglas son estudiadas por varios autores, incluyendo Scheffé [98d], Bennett y Franklin [9], Cornfield y Tukey [34] y Searle [99a, b]. Mediante el examen de los cuadrados medios esperados puede desarrollarse el estadístico apropiado para probar hipótesis acerca de cualquier parámetro del modelo. El estadístico de prueba es el cociente de los cuadrados medios que se elige, de tal modo que el valor esperado del cuadrado medio del **numerador** difiere del valor esperado del cuadrado medio del **denominador** únicamente por el componente de la varianza o el factor fijo en el que se tiene interés.

Siempre es posible determinar los cuadrados medios esperados de cualquier modelo como se hizo en el capítulo 3, es decir, mediante la aplicación directa del operador valor esperado. Este método de **fuerza bruta,** como suele llamársele, puede ser muy laborioso. Las reglas que se presentan a continuación producen siempre los cuadrados medios esperados sin recurrir al enfoque de fuerza bruta y, con la práctica, su uso se vuelve relativamente simple. Cuando se aplican a un modelo mixto, estas reglas producen cuadrados medios esperados que son consistentes con los supuestos del **modelo mixto restringido** de la sección 12-3. Las reglas se ilustran utilizando el modelo factorial de efectos fijos con dos factores.

Regla 1. El término del error del modelo, $\varepsilon_{ij...m}$, se escribe como $\varepsilon_{(ij...)m}$, donde el subíndice m denota el subíndice de la réplica. Para el modelo con dos factores, esta regla implica que ε_{ijk} se convierte en $\varepsilon_{(ij)k}$.

Regla 2. Además de una media global (μ) y un término del error [$\varepsilon_{(ij...)m}$], el modelo contiene todos los efectos principales y las interacciones cuya existencia supone el experimentador. Si existen todas las interacciones posibles entre los k factores, entonces hay $\binom{k}{2}$ interacciones de dos factores, $\binom{k}{3}$ interacciones de tres factores, ..., 1 interacción de k factores. Si uno de los factores de un término aparece entre paréntesis, entonces no hay interacción entre ese factor y los demás factores de ese término.

Regla 3. Para cada término del modelo, los subíndices se dividen en tres clases: *a*) vivos: aquellos que están presentes en el término y no están entre paréntesis; *b*) muertos: aquellos que están presentes en el término y están entre paréntesis; y *c*) ausentes: aquellos subíndices que están presentes en el modelo pero no en ese término particular.

Por lo tanto, en $(\tau\beta)_{ij}$, i y j son subíndices vivos y k es un subíndice ausente, y en $\varepsilon_{(ij)k}$, k es un subíndice vivo, mientras que i y j son subíndices muertos.

[2] Los diseños anidados se estudian en el capítulo 13.

Regla 4. Grados de libertad. El número de grados de libertad de cualquier término del modelo es el producto del número de niveles asociados con cada subíndice muerto y el número de niveles asociados con cada subíndice vivo menos 1.

Por ejemplo, el número de grados de libertad asociados con $(\tau\beta)_{ij}$ es $(a-1)(b-1)$, y el número de grados de libertad asociados con $\varepsilon_{(ij)k}$ es $ab(n-1)$.

Regla 5. Cada término del modelo tiene asociado con él un componente de la varianza (efecto aleatorio) o bien un factor fijo (efecto fijo). Si una interacción contiene al menos un efecto aleatorio, la interacción completa se considera aleatoria. Un componente de la varianza tiene letras griegas como subíndices para identificar el efecto aleatorio particular. Por lo tanto, en un modelo mixto de dos factores con el factor A fijo y el factor B aleatorio, el componente de la varianza de B es σ_β^2, y el componente de la varianza de AB es $\sigma_{\tau\beta}^2$. Un efecto fijo se representa siempre por la suma de cuadrados de los componentes del modelo asociados con ese factor dividida por sus grados de libertad. En el ejemplo tratado aquí, el efecto de A es

$$\frac{\sum_{i=1}^{a} \tau_i^2}{a-1}$$

Regla 6. Cuadrados medios esperados. Para obtener los cuadrados medios esperados, se elabora la tabla siguiente. Hay un renglón para cada componente (cuadrado medio) del modelo y una columna para cada subíndice. Arriba de cada subíndice se escribe el número de niveles del factor asociados con ese subíndice y si el factor es fijo (F) o aleatorio (R). Las réplicas siempre se consideran aleatorias.

a) En cada renglón se escribe 1 si uno de los subíndices muertos en el componente del renglón coincide con el subíndice de la columna:

	F	F	R
	a	b	n
Factor	i	j	k
τ_i			
β_j			
$(\tau\beta)_{ij}$			
$\varepsilon_{(ij)k}$	1	1	

b) En cada renglón, si cualquiera de los subíndices del componente del mismo coincide con el subíndice de la columna, se escribe 0 si el encabezado de la columna es un factor fijo y 1 si es un factor aleatorio:

	F	F	R
	a	b	n
Factor	i	j	k
τ_i	0		
β_j		0	
$(\tau\beta)_{ij}$	0	0	
$\varepsilon_{(ij)k}$	1	1	1

c) En las posiciones del renglón que quedan vacías se escribe el número de niveles que aparecen indicados arriba del encabezado de la columna:

	F	F	R
	a	b	n
Factor	i	j	k
τ_i	0	b	n
β_j	a	0	n
$(\tau\beta)_{ij}$	0	0	n
$\varepsilon_{(ij)k}$	1	1	1

d) Para obtener los cuadrados medios esperados de cualquier componente del modelo, primero se cubren todas las columnas cuyos encabezados sean subíndices vivos de ese componente. Después, en cada renglón que contiene *al menos* los *mismos* subíndices que los del componente bajo consideración, se toma el producto de los números visibles y se multiplica por el factor fijo o aleatorio apropiado de la regla 1. La suma de estas cantidades es el cuadrado medio esperado del componente del modelo bajo consideración. Por ejemplo, para encontrar $E(MS_A)$, se cubre la columna i. El producto de los números visibles en los renglones que contienen al menos el subíndice i son bn (renglón 1), 0 (renglón 3) y 1 (renglón 4). Observe que i no está presente en el renglón 2. Por lo tanto, el cuadrado medio esperado es

$$E(MS_A) = \sigma^2 + \frac{bn\sum_{i=1}^{a} \tau_i^2}{a-1}$$

En la tabla 12-9 se presenta la tabla completa de los cuadrados medios esperados para este diseño. En las tablas 12-10 y 12-11 se muestran las derivaciones de los cuadrados medios esperados para los modelos con dos factores, aleatorio y mixto, respectivamente. Observe que se ha supuesto la versión restringida del modelo mixto para producir los cuadrados medios esperados. En el ejemplo siguiente se considera un diseño factorial con tres factores.

EJEMPLO 12-6 ··

Considere un experimento factorial de tres factores con a niveles del factor A, b niveles del factor B, c niveles del factor C y n réplicas. El análisis de este diseño, suponiendo que todos los factores son de efectos

Tabla 12-9 Derivación de los cuadrados medios esperados, modelo de efectos fijos con dos factores

	F	F	R	
	a	b	n	Cuadrado
Factor	i	j	k	medio esperado
τ_i	0	b	n	$\sigma^2 + \frac{bn\Sigma\tau_i^2}{a-1}$
β_j	a	0	n	$\sigma^2 + \frac{an\Sigma\beta_j^2}{b-1}$
$(\tau\beta)_{ij}$	0	0	n	$\sigma^2 + \frac{n\Sigma\Sigma(\tau\beta)_{ij}^2}{(a-1)(b-1)}$
$\varepsilon_{(ij)k}$	1	1	1	σ^2

Tabla 12-10 Derivación de los cuadrados medios esperados, modelo de efectos aleatorios con dos factores

Factor	R a i	R b j	R n k	Cuadrado medio esperado
τ_i	1	b	n	$\sigma^2 + n\sigma_{\tau\beta}^2 + bn\sigma_\tau^2$
β_j	a	1	n	$\sigma^2 + n\sigma_{\tau\beta}^2 + an\sigma_\beta^2$
$(\tau\beta)_{ij}$	1	1	n	$\sigma^2 + n\sigma_{\tau\beta}^2$
$\varepsilon_{(ij)k}$	1	1	1	σ^2

Tabla 12-11 Derivación de los cuadrados medios esperados, modelo mixto con dos factores

Factor	F a i	R b j	R n k	Cuadrado medio esperado
τ_i	0	b	n	$\sigma^2 + n\sigma_{\tau\beta}^2 + \frac{bn\Sigma\tau_i^2}{a-1}$
β_j	a	1	n	$\sigma^2 + an\sigma_\beta^2$
$(\tau\beta)_{ij}$	0	1	n	$\sigma^2 + n\sigma_{\tau\beta}^2$
$\varepsilon_{(ij)k}$	1	1	1	σ^2

Tabla 12-12 Derivación de los cuadrados medios esperados, modelo de efectos aleatorios con tres factores

Factor	R a i	R b j	R c k	R n l	Cuadrado medio esperado
τ_i	1	b	c	n	$\sigma^2 + cn\sigma_{\tau\beta}^2 + bn\sigma_{\tau\gamma}^2 + n\sigma_{\tau\beta\gamma}^2 + bcn\sigma_\tau^2$
β_j	a	1	c	n	$\sigma^2 + cn\sigma_{\tau\beta}^2 + an\sigma_{\beta\gamma}^2 + n\sigma_{\tau\beta\gamma}^2 + acn\sigma_\beta^2$
γ_k	a	b	1	n	$\sigma^2 + bn\sigma_{\tau\gamma}^2 + an\sigma_{\beta\gamma}^2 + n\sigma_{\tau\beta\gamma}^2 + abn\sigma_\gamma^2$
$(\tau\beta)_{ij}$	1	1	c	n	$\sigma^2 + n\sigma_{\tau\beta\gamma}^2 + cn\sigma_{\tau\beta}^2$
$(\tau\gamma)_{ik}$	1	b	1	n	$\sigma^2 + n\sigma_{\tau\beta\gamma}^2 + bn\sigma_{\tau\gamma}^2$
$(\beta\gamma)_{jk}$	a	1	1	n	$\sigma^2 + n\sigma_{\tau\beta\gamma}^2 + an\sigma_{\beta\gamma}^2$
$(\tau\beta\gamma)_{ijk}$	1	1	1	n	$\sigma^2 + n\sigma_{\tau\beta\gamma}^2$
ε_{ijkl}	1	1	1	1	σ^2

fijos, se presenta en la sección 5-4. Ahora se determinan los cuadrados medios esperados suponiendo que todos los factores son *aleatorios*. El modelo estadístico apropiado es

$$y_{ijkl} = \mu + \tau_i + \beta_j + \gamma_k + (\tau\beta)_{ij} + (\tau\gamma)_{ik} + (\beta\gamma)_{jk} + (\tau\beta\gamma)_{ijk} + \varepsilon_{ijkl}$$

Utilizando las reglas descritas antes, en la tabla 12-12 se derivan los cuadrados medios esperados.

Se observa, al examinar los cuadrados medios esperados de la tabla 12-12, que si A, B y C son factores aleatorios, entonces no existe ninguna prueba exacta para los efectos principales. Es decir, si se quiere probar la hipótesis $\sigma_\tau^2 = 0$, no es posible formar un cociente de dos cuadrados medios esperados tal que el único término del numerador que no está en el denominador sea $bcn\sigma_\tau^2$. El mismo fenómeno ocurre para los efectos principales de B y C. Observe que efectivamente existen las pruebas apropiadas para las interacciones de dos y tres factores. Sin embargo, es posible que las pruebas de los efectos principales sean de importancia básica para el experimentador. Por lo tanto, ¿cómo deberán probarse los efectos principales? Este problema se considera en la siguiente sección.

12-6 PRUEBAS F APROXIMADAS

Es frecuente que en experimentos factoriales con tres o más factores incluidos en un modelo aleatorio o mixto, así como en otros diseños más complejos, no exista un estadístico de prueba exacto para ciertos efectos de los modelos. Una posible solución a este dilema es suponer que ciertas interacciones son insignificantes. Para ilustrar, si fuera razonable suponer que todas las interacciones de dos factores del ejemplo 12-6 son insignificantes, entonces podría hacerse $\sigma_{\tau\beta}^2 = \sigma_{\tau\gamma}^2 = \sigma_{\beta\gamma}^2 = 0$, y sería posible conducir pruebas de los efectos principales.

Aun cuando parece tratarse de una posibilidad atractiva, es necesario señalar que debe haber algo en la naturaleza del proceso —o algún conocimiento previo sólido— que permita suponer que una o más de las interacciones son insignificantes. En general, no es sencillo establecer este supuesto, y tampoco deberá hacerse a la ligera. No deberán eliminarse ciertas interacciones del modelo sin evidencia concluyente de que es apropiado hacerlo. Un procedimiento defendido por algunos experimentadores es probar primero las interacciones, después fijar en cero aquellas interacciones que se hayan encontrado no significativas, para después suponer que estas interacciones son cero cuando se prueben otros efectos en el mismo experimento. Aun cuando en ocasiones se aplica en la práctica, este procedimiento puede ser riesgoso, ya que cualquier decisión respecto a una interacción está sujeta tanto al error tipo I como al error tipo II.

Una variante de esta idea es **agrupar** ciertos cuadrados medios en el análisis de varianza para obtener una estimación del error con más grados de libertad. Por ejemplo, suponga que en el ejemplo 12-6 no fue significativo el estadístico de prueba $F_0 = MS_{ABC}/MS_E$. Por lo tanto, $H_0{:}\sigma_{\tau\beta\gamma}^2 = 0$ no se rechaza, y *tanto* MS_{ABC} *como* MS_E estiman la varianza del error σ^2. El experimentador podría considerar la agrupación o combinación de MS_{ABC} y MS_E de acuerdo con

$$MS_{E'} = \frac{abc(n-1)MS_E + (a-1)(b-1)(c-1)MS_{ABC}}{abc(n-1)+(a-1)(b-1)(c-1)}$$

de tal modo que $E(MS_{E'}) = \sigma^2$. Observe que $MS_{E'}$ tiene $abc(n-1) + (a-1)(b-1)(c-1)$ grados de libertad, en comparación con los $abc(n-1)$ grados de libertad del MS_E original.

El riesgo de agrupar es que puede incurrirse en un error tipo II y combinar con el error el cuadrado medio de un factor que en realidad *es* significativo, obteniéndose así un nuevo cuadrado medio residual

($MS_{E'}$) que es muy grande. Esto hará que sea más difícil detectar otros efectos significativos. Por otra parte, si el cuadrado medio del error original tiene un número muy pequeño de grados de libertad (por ejemplo, menos de seis), el experimentador quizá tenga mucho que ganar al hacer la agrupación, ya que podría conseguirse así un incremento potencialmente considerable de la precisión de pruebas posteriores. Un procedimiento razonablemente práctico es el siguiente. Si el cuadrado medio del error original tiene seis o más grados de libertad, no hacer la agrupación. Si el cuadrado medio del error original tiene menos de seis grados de libertad, hacer la agrupación sólo si el estadístico F del cuadrado medio que se agrupará no es significativo para un valor grande de α, tal como $\alpha = 0.25$.

Cuando no es posible suponer que ciertas interacciones son insignificantes y sigue siendo necesario hacer inferencias acerca de los efectos para los que no existen pruebas exactas, puede emplearse un procedimiento atribuido a Satterthwaite [97]. El método de Satterthwaite utiliza **combinaciones lineales de cuadrados medios,** por ejemplo,

$$MS' = MS_r + \cdots + MS_s \tag{12-31}$$

y

$$MS'' = MS_u + \cdots + MS_v \tag{12-32}$$

donde los cuadrados medios de las ecuaciones 12-31 y 12-32 se seleccionan de tal modo que $E(MS') - E(MS'')$ sea igual a un múltiplo del efecto (el parámetro del modelo o el componente de la varianza) considerado en la hipótesis nula. Entonces el estadístico de prueba sería

$$F = \frac{MS'}{MS''} \tag{12-33}$$

que se distribuye aproximadamente como $F_{p,\,q}$, donde

$$p = \frac{(MS_r + \cdots + MS_s)^2}{MS_r^2 / f_r + \cdots + MS_s^2 / f_s} \tag{12-34}$$

y

$$q = \frac{(MS_u + \cdots + MS_v)^2}{MS_u^2 / f_u + \cdots + MS_v^2 / f_v} \tag{12-35}$$

En p y q, f_i es el número de grados de libertad asociados con el cuadrado medio MS_i. No existe la seguridad de que p y q sean enteros, por lo que puede ser necesario hacer una interpolación en las tablas de la distribución F. Por ejemplo, en el modelo de efectos aleatorios con tres factores (tabla 12-12), es relativamente sencillo ver que un estadístico de prueba apropiado para $H_0{:}\sigma_\tau^2 = 0$ sería $F = MS'/MS''$, con

$$MS' = MS_A + MS_{ABC}$$

y

$$MS'' = MS_{AB} + MS_{AC}$$

Los grados de libertad de F se calcularían con las ecuaciones 12-34 y 12-35.

La teoría subyacente de esta prueba es que tanto el numerador como el denominador del estadístico de prueba (ecuación 12-33) se distribuyen aproximadamente como múltiplos de variables aleatorias

ji-cuadrada, y puesto que no aparece ningún cuadrado medio en el numerador o el denominador de la ecuación 12-33, el numerador y el denominador son independientes. Por lo tanto, en la ecuación 12-33, F se distribuye aproximadamente como $F_{p,q}$. Satterthwaite hace hincapié en que deberá prestarse atención al aplicar el procedimiento cuando algunos de los cuadrados medios de MS' y MS'' aparezcan con signos negativos. Gaylor y Hopper [48] reportan que si $MS' = MS_1 - MS_2$, entonces la aproximación de Satterthwaite tiene una validez razonable si

$$\frac{MS_1}{MS_2} > F_{0.025, f_2, f_1} \times F_{0.50, f_2, f_2}$$

y si $f_1 \leq 100$ y $f_2 \geq f_1/2$.

EJEMPLO 12-7

Se estudia la caída de la presión medida en una válvula de expansión de una turbina. El ingeniero de diseño considera que las variables importantes que influyen en las lecturas de la caída de la presión son la temperatura del gas en la admisión (A), el operador (B) y el manómetro específico que utiliza el operador (C). Estos tres factores se incluyen en un diseño factorial, con la temperatura del gas fija y el operador y el manómetro aleatorios. En la tabla 12-13 se muestran los datos codificados de dos réplicas. El modelo lineal de este diseño es

$$y_{ijkl} = \mu + \tau_i + \beta_j + \gamma_k + (\tau\beta)_{ij} + (\tau\gamma)_{ik} + (\beta\gamma)_{jk} + (\tau\beta\gamma)_{ijk} + \varepsilon_{ijkl}$$

donde τ_i es el efecto de la temperatura del gas (A), β_j es el efecto del operador (B) y γ_k es el efecto del manómetro (C).

El análisis de varianza se muestra en la tabla 12-14. Se ha agregado la columna titulada "Cuadrados medios esperados" a esta tabla, y las entradas de esta columna se derivan por los métodos estudiados en la sección 12-5. Por la columna Cuadrados medios esperados, se observa que existen pruebas exactas para todos los efectos esperados salvo el efecto principal A. En la tabla 12-14 se muestran los resultados de estas pruebas. Para probar el efecto de la temperatura del gas, o H_0:$\tau_i = 0$, podría usarse el estadístico

$$F = \frac{MS'}{MS''}$$

Tabla 12-13 Datos codificados de la caída de la presión para el experimento de la turbina

	Temperatura del gas (A)											
	60°F				75°F				90°F			
Manómetro	Operador (B)				Operador (B)				Operador (B)			
(C)	1	2	3	4	1	2	3	4	1	2	3	4
1	−2	0	−1	4	14	6	1	−7	−8	−2	−1	−2
	−3	−9	−8	4	14	0	2	6	−8	20	−2	1
2	−6	−5	−8	−3	22	8	6	−5	−8	1	−9	−8
	4	−1	−2	−7	24	6	2	2	3	−7	−8	3
3	−1	−4	0	−2	20	2	3	−5	−2	−1	−4	1
	−2	−8	−7	4	16	0	0	−1	−1	−2	−7	3

Tabla 12-14 Análisis de varianza de los datos de la caída de la presión

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrados medios esperados	Cuadrado medio	F_0	Valor P
Temperatura, A	1023.36	2	$\sigma^2 + bn\sigma_{\tau\gamma}^2 + cn\sigma_{\tau\beta}^2 + n\sigma_{\tau\beta\gamma}^2 + \frac{bcn\Sigma\tau_i^2}{a-1}$	511.68	2.22	0.17
Operador, B	423.82	3	$\sigma^2 + an\sigma_{\beta\gamma}^2 + acn\sigma_{\beta}^2$	141.27	4.05	0.07
Manómetro, C	7.19	2	$\sigma^2 + an\sigma_{\beta\gamma}^2 + abn\sigma_{\gamma}^2$	3.60	0.10	0.90
AB	1211.97	6	$\sigma^2 + n\sigma_{\tau\beta\gamma}^2 + cn\sigma_{\tau\beta}^2$	202.00	14.59	<0.01
AC	137.89	4	$\sigma^2 + n\sigma_{\tau\beta\gamma}^2 + bn\sigma_{\tau\gamma}^2$	34.47	2.49	0.10
BC	209.47	6	$\sigma^2 + an\sigma_{\beta\gamma}^2$	34.91	1.63	0.17
ABC	166.11	12	$\sigma^2 + n\sigma_{\tau\beta\gamma}^2$	13.84	0.65	0.79
Error	770.50	36	σ^2	21.40		
Total	3950.32	71				

donde

$$MS' = MS_A + MS_{ABC}$$

y

$$MS'' = MS_{AB} + MS_{AC}$$

ya que

$$E(MS') - E(MS'') = \frac{bcn\Sigma\tau_i^2}{a-1}$$

Para determinar el estadístico de prueba para $H_0{:}\tau_i = 0$, se calculan

$$\begin{aligned} MS' &= MS_A + MS_{ABC} \\ &= 511.68+13.84 = 525.52 \\ MS'' &= MS_{AB} + MS_{AC} \\ &= 202.00+34.47 = 236.47 \end{aligned}$$

y

$$F = \frac{MS'}{MS''} = \frac{525.52}{236.47} = 2.22$$

Los grados de libertad de este estadístico se encuentran con las ecuaciones 12-34 y 12-35 de la siguiente manera:

$$\begin{aligned} p &= \frac{(MS_A + MS_{ABC})^2}{MS_A^2/2 + MS_{ABC}^2/12} \\ &= \frac{(525.52)^2}{(511.68)^2/2+(13.84)^2/12} = 2.11 \simeq 2 \end{aligned}$$

y

$$\begin{aligned} q &= \frac{(MS_{AB} + MS_{AC})^2}{MS_{AB}^2/6 + MS_{AC}^2/4} \\ &= \frac{(236.47)^2}{(202.00)^2/6+(34.47)^2/4} = 7.88 \simeq 8 \end{aligned}$$

Al comparar $F = 2.22$ con $F_{0.05,2,8} = 4.46$, no puede rechazarse H_0. El valor P aproximado es $P = 0.17$.

La interacción AB, o temperatura-operador, es grande, y hay ciertos indicios de una interacción AC o temperatura-manómetro. El análisis gráfico de las interacciones AB y AC, ilustrado en la figura 12-2, indica que el efecto de la temperatura puede ser grande cuando se usan el operador 1 y el manómetro 3. Por lo tanto, parece posible que los efectos principales de la temperatura y el operador estén enmascarados por la interacción AB grande.

....

En la tabla 12-15 se presenta la salida de la rutina Balanced ANOVA (análisis de varianza balanceado) de Minitab para el experimento del ejemplo 12-7. Se ha especificado el **modelo restringido.** Q[1] representa el efecto fijo de la presión del gas. Observe que las entradas de la tabla del análisis de varianza

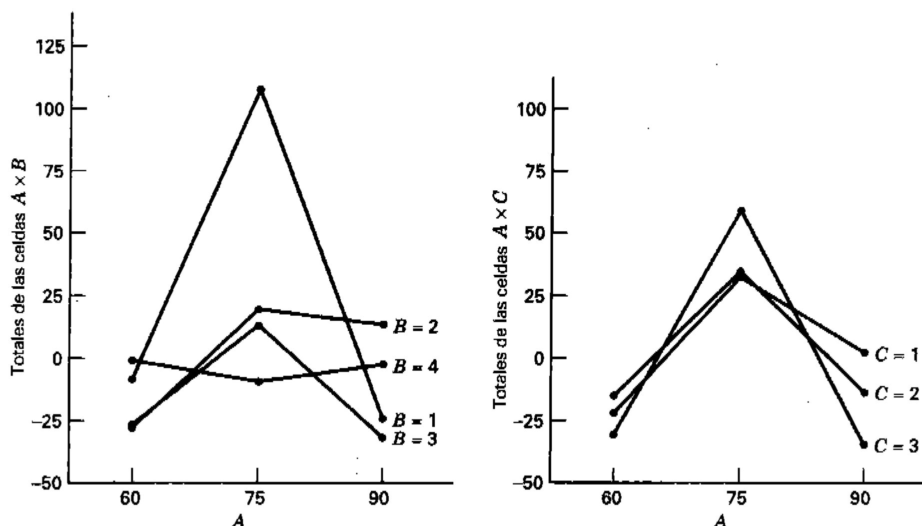


Figura 12-2 Interacciones en el experimento de la caída de la presión.

concuerdan en general con las de la tabla 12-14, salvo por la prueba F de la temperatura del gas (factor A). Minitab indica que la prueba no es exacta (lo que se ve por los cuadrados medios esperados). La Prueba Sintetizada construida por Minitab es en realidad el procedimiento de Satterthwaite, pero usa un estadístico de prueba diferente del que se utilizó aquí. Observe que, por la salida de Minitab, el cuadrado medio del error para probar el factor A es

$$(4)+(5)-7 = MS_{AB} + MS_{AC} - MS_{ABC}$$

cuyo valor esperado es

$$\begin{aligned}E[(4)+(5)-(7)] &= \sigma^2 + n\sigma^2_{\tau\beta\gamma} + cn\sigma^2_{\tau\beta} + \sigma^2 + n\sigma^2_{\tau\beta\gamma} \\ &\quad + bn\sigma^2_{\tau\gamma} - (\sigma^2 + n\sigma^2_{\tau\beta\gamma}) \\ &= \sigma^2 + n\sigma^2_{\tau\beta\gamma} + cn\sigma^2_{\tau\beta} + bn\sigma^2_{\tau\gamma}\end{aligned}$$

que es un cuadrado medio del error apropiado para probar el efecto promedio de A. Esto es una muy buena ilustración de que puede haber más de una manera de construir los cuadrados medios sintéticos usados en el procedimiento de Satterthwaite. Sin embargo, se preferiría en general la combinación lineal de los cuadrados medios que se seleccionaron, en lugar de la que eligió Minitab, ya que no incluye ningún cuadrado medio con signo negativo en las combinaciones lineales.

El análisis del ejemplo 12-7, suponiendo el modelo no restringido, se presenta en la tabla 12-16. La diferencia principal con el modelo restringido es que ahora los valores esperados de los cuadrados medios de los tres efectos principales son tales que no existe ninguna prueba exacta. En el modelo restringido, los dos efectos aleatorios promedio podrían probarse contra su interacción, pero ahora el cuadrado medio esperado de B incluye a $\sigma^2_{\tau\beta\gamma}$ y a $\sigma^2_{\tau\beta}$, y el cuadrado medio esperado de C incluye a $\sigma^2_{\tau\beta\gamma}$ y a $\sigma^2_{\tau\gamma}$. De nueva cuenta, Minitab construye cuadrados medios sintéticos y prueba estos efectos con el procedimiento de Satterthwaite. Las conclusiones generales no son radicalmente diferentes del análisis del modelo restringido, además del cambio grande en la estimación del componente de la varianza del operador. El modelo no restringido produce una estimación negativa de σ^2_β. Puesto que el factor manómetro no es significativo en ninguno de los dos análisis, es posible que sea pertinente alguna reducción del modelo.

Tabla 12-15 Análiss de varianza balanceado de Minitab (Balanced ANOVA) del ejemplo 12-7, modelo restringido

Análisis de varianza (diseños balanceados)

```
Factor      Type Levels Values
GasT       fixed      3     60     75     90
Operator random       4      1      2      3      4
Gauge     random      3      1      2      3

Analysis of Variance for Drop

Source                 DF        SS        MS        F       P
GasT                    2   1023.36    511.68     2.30   0.171 x
Operator                3    423.82    141.27     4.05   0.069
Gauge                   2      7.19      3.60     0.10   0.904
GasT*Operator           6   1211.97    202.00    14.59   0.000
GasT*Gauge              4    137.89     34.47     2.49   0.099
Operator*Gauge          6    209.47     34.91     1.63   0.167
GasT*Operator*Gauge    12    166.11     13.84     0.65   0.788
Error                  36    770.50     21.40
Total                  71   3950.32

x Not an exact F-test.

Source                 Variance Error Expected Mean Square for Each Term
                      component  term (using restricted model)
 1 GasT                            *   (8) + 2(7) + 8(5) + 6(4) + 24Q[1]
 2 Operator               5.909    6   (8) + 6(6) + 18(2)
 3 Gauge                 -1.305    6   (8) + 6(6) + 24(3)
 4 GasT*Operator         31.359    7   (8) + 2(7) + 6(4)
 5 GasT*Gauge             2.579    7   (8) + 2(7) + 8(5)
 6 Operator*Gauge         2.252    8   (8) + 6(6)
 7 GasT*Operator*Gauge   -3.780    8   (8) + 2(7)
 8 Error                 21.403        (8)

* Synthesized Test.

Error Terms for Synthesized Tests

Source                 Error DF  Error MS  Synthesis of Error MS
 1 GasT                    6.97    222.63   (4) + (5) - (7)
```

Tabla 12-16 Análisis de varianza balanceado de Minitab (Balanced ANOVA) del ejemplo 12-7, modelo no restringido

Análisis de varianza (diseños balanceados)

```
Factor      Type Levels Values
GasT       fixed      3     60     75     90
Operator  random      4      1      2      3      4
Gauge     random      3      1      2      3

Analysis of Variance for Drop

Source                  DF        SS        MS       F       P
GasT                     2   1023.36    511.68    2.30   0.171 x
Operator                 3    423.82    141.27    0.63   0.616 x
Gauge                    2      7.19      3.60    0.06   0.938 x
GasT*Operator            6   1211.97    202.00   14.59   0.000
GasT*Gauge               4    137.89     34.47    2.49   0.099
Operator*Gauge           6    209.47     34.91    2.52   0.081
GasT*Operator*Gauge     12    166.11     13.84    0.65   0.788
Error                   36    770.50     21.40
Total                   71   3950.32

x Not an exact F-test.

Source                 Variance Error Expected Mean Square for Each Term
                      component term (using unrestricted model)
 1 GasT                           *   (8) + 2(7) + 8(5) + 6(4) + Q[1]
 2 Operator              -4.544   *   (8) + 2(7) + 6(6) + 6(4) + 18(2)
 3 Gauge                 -2.164   *   (8) + 2(7) + 6(6) + 8(5) + 24(3)
 4 GasT*Operator         31.359   7   (8) + 2(7) + 6(4)
 5 GasT*Gauge             2.579   7   (8) + 2(7) + 8(5)
 6 Operator*Gauge         3.512   7   (8) + 2(7) + 6(6)
 7 GasT*Operator*Gauge   -3.780   8   (8) + 2(7)
 8 Error                 21.403       (8)

* Synthesized Test.

Error Terms for Synthesized Tests

Source                 Error DF  Error MS  Synthesis of Error MS
 1 GasT                    6.97    222.63   (4) + (5) - (7)
 2 Operator                7.09    223.06   (4) + (6) - (7)
 3 Gauge                   5.98     55.54   (5) + (6) - (7)
```

12-7 ALGUNOS TEMAS ADICIONALES SOBRE LA ESTIMACIÓN DE LOS COMPONENTES DE LA VARIANZA

Como se señaló anteriormente, la estimación de los componentes de la varianza en un modelo aleatorio o mixto reviste con frecuencia considerable importancia para el experimentador. En esta sección se presentan algunos resultados y técnicas adicionales que son útiles para estimar los componentes de la varianza. La atención se centra en los procedimientos para encontrar intervalos de confianza para los componentes de la varianza, y se ilustra asimismo cómo encontrar estimaciones de máxima verosimilitud de los componentes de la varianza. El método de máxima verosimilitud puede ser una alternativa útil cuando el método del análisis de varianza produce estimaciones negativas.

12-7.1 Intervalos de confianza aproximados para los componentes de la varianza

Cuando se introdujo el modelo de efectos aleatorios en la sección 12-1, se presentaron intervalos de confianza exactos de $100(1-\alpha)$ por ciento para σ^2 y para otras funciones de los componentes de la varianza en ese diseño experimental simple. Siempre es posible encontrar un intervalo de confianza exacto para cualquier función de los componentes de la varianza que es el valor esperado de uno de los cuadrados medios del análisis de varianza. Por ejemplo, considere el cuadrado medio del error. Puesto que $E(MS_E) = \sigma^2$, siempre es posible encontrar un intervalo de confianza exacto para σ^2, ya que la cantidad

$$f_E MS_E / \sigma^2 = f_E \hat{\sigma}^2 / \sigma^2$$

tiene una distribución ji-cuadrada con f_E grados de libertad. El intervalo de confianza exacto de $100(1-\alpha)$ por ciento es

$$\frac{f_E MS_E}{\chi^2_{\alpha/2,f_E}} \leq \sigma^2 \leq \frac{f_E MS_E}{\chi^2_{1-\alpha/2,f_E}} \tag{12-36}$$

Desafortunadamente, en experimentos más complejos en los que intervienen varios factores del diseño no es posible por lo general encontrar intervalos de confianza exactos para los componentes de la varianza de interés, ya que estas varianzas no son el valor esperado de un solo cuadrado medio del análisis de varianza. Sin embargo, los conceptos fundamentales de las "pseudo" pruebas F aproximadas de Satterthwaite, introducidos en la sección 12-6, pueden emplearse para construir intervalos de confianza aproximados de los componentes de la varianza para los que no se cuenta con ningún intervalo de confianza exacto.

Recuerde que el método de Satterthwaite utiliza dos combinaciones lineales de cuadrados medios

$$MS' = MS_r + \cdots + MS_s$$

y

$$MS'' = MS_u + \cdots + MS_v$$

con el estadístico de prueba

$$F = \frac{MS'}{MS''}$$

que tiene una distribución F aproximada. Al utilizar los grados de libertad apropiados para MS' y MS'', definidos en las ecuaciones 12-34 y 12-35, este estadístico F puede usarse en una prueba de significación aproximada del parámetro o del componente de la varianza de interés.

Para probar la significación de un componente de la varianza, por ejemplo σ_0^2, las dos combinaciones lineales, MS' y MS'', se eligen de tal modo que la diferencia en sus valores esperados sea igual a un múltiplo del componente, por ejemplo

$$E(MS') - E(MS'') = k\sigma_0^2$$

o

$$\sigma_0^2 = \frac{E(MS') - E(MS'')}{k}. \tag{12-37}$$

La ecuación 12-37 proporciona una base para una estimación puntual de σ_0^2:

$$\begin{aligned}\hat{\sigma}_0^2 &= \frac{MS' - MS''}{k} \\ &= \frac{1}{k}MS_r + \cdots + \frac{1}{k}MS_s - \frac{1}{k}MS_u - \cdots \frac{1}{k}MS_v\end{aligned} \tag{12-38}$$

Los cuadrados medios (MS_i) de la ecuación 12-38 son independientes, donde las $f_iMS_i/\sigma_i^2 = SS_i/\sigma_i^2$ tienen distribuciones ji-cuadrada con f_i grados de libertad. La estimación del componente de la varianza, $\hat{\sigma}_0^2$, es una combinación lineal de múltiplos de los cuadrados medios, y $r\hat{\sigma}_0^2/\sigma_0^2$ sigue una distribución ji-cuadrada aproximada con r grados de libertad, donde

$$\begin{aligned}r &\simeq \frac{(\hat{\sigma}_0^2)^2}{\sum_{i=1}^{m} \frac{1}{k^2}\frac{MS_i^2}{f_i}} \\ &= \frac{(MS_r + \cdots + MS_s - MS_u - \cdots - MS_v)^2}{\frac{MS_r^2}{f_r} + \cdots + \frac{MS_s^2}{f_s} + \frac{MS_u^2}{f_u} + \cdots + \frac{MS_v^2}{f_v}}\end{aligned} \tag{12-39}$$

Este resultado sólo puede usarse si $\hat{\sigma}_0^2 > 0$. Como r no será un entero en la mayoría de los casos, por lo general será necesario hacer una interpolación de las tablas ji-cuadrada. Graybill [50] establece un resultado general para r.

Ahora bien, puesto que $r\hat{\sigma}_0^2/\sigma^2$ tiene una distribución ji-cuadrada aproximada con r grados de libertad,

$$P\left\{\chi^2_{1-\alpha/2,r} \le \frac{r\hat{\sigma}_0^2}{\sigma_0^2} \le \chi^2_{\alpha/2,r}\right\} = 1-\alpha$$

y

$$P\left\{\frac{r\hat{\sigma}_0^2}{\chi^2_{\alpha/2,r}} \le \sigma_0^2 \le \frac{r\hat{\sigma}_0^2}{\chi^2_{1-\alpha/2,r}}\right\} = 1-\alpha$$

Por lo tanto, un intervalo de confianza aproximado de $100(1-\alpha)$ por ciento para σ_0^2 es

$$\frac{r\hat{\sigma}_0^2}{\chi^2_{\alpha/2,r}} \le \sigma_0^2 \le \frac{r\hat{\sigma}_0^2}{\chi^2_{1-\alpha/2,r}} \tag{12-40}$$

EJEMPLO 12-8 ··

Para ilustrar este procedimiento, considere nuevamente el experimento del ejemplo 12-7, donde se usó un modelo mixto con tres factores en un estudio de la caída de la presión en una válvula de expansión de una turbina. El modelo es

$$y_{ijkl} = \mu + \tau_i + \beta_j + \gamma_k + (\tau\beta)_{ij} + (\tau\gamma)_{ik} + (\beta\gamma)_{jk} + (\tau\beta\gamma)_{ijk} + \varepsilon_{ijkl}$$

donde τ_i es un efecto fijo y todos los demás efectos son aleatorios. Se encontrará un intervalo de confianza aproximado para $\sigma^2_{\tau\beta}$. Al utilizar los cuadrados medios esperados de la tabla 12-14, se observa que la diferencia en los valores esperados de los cuadrados medios para el efecto de la interacción de dos factores AB y el efecto de interacción de tres factores ABC es un múltiplo del componente de la varianza de interés, $\sigma^2_{\tau\beta}$.

$$E(MS_{AB}) - E(MS_{ABC}) = \sigma^2 + n\sigma^2_{\tau\beta\gamma} + cn\sigma^2_{\tau\beta} - (\sigma^2 + n\sigma^2_{\tau\beta\gamma})$$
$$= cn\sigma^2_{\tau\beta}$$

Por lo tanto, la estimación puntual de $\sigma^2_{\tau\beta}$ es

$$\hat{\sigma}^2_{\tau\beta} = \frac{MS_{AB} - MS_{ABC}}{cn} = \frac{134.91 - 19.26}{(3)(2)} = 19.28$$

y

$$r = \frac{(MS_{AB} - MS_{ABC})^2}{\dfrac{MS^2_{AB}}{(a-1)(b-1)} + \dfrac{MS^2_{ABC}}{(a-1)(b-1)(c-1)}} = \frac{(134.91 - 19.26)^2}{\dfrac{(134.91)^2}{(2)(3)} + \dfrac{(19.26)^2}{(2)(3)(2)}} = 4.36$$

El intervalo de confianza aproximado de 95% para $\sigma^2_{\tau\beta}$ se encuentra entonces con la ecuación 12-40 de la siguiente manera:

$$\frac{r\hat{\sigma}^2_{\tau\beta}}{\chi^2_{0.025,r}} \le \sigma^2_{\tau\beta} \le \frac{r\hat{\sigma}^2_{\tau\beta}}{\chi^2_{0.975,r}}$$
$$\frac{(4.36)(19.28)}{11.58} \le \sigma^2_{\tau\beta} \le \frac{(4.36)(19.28)}{0.61}$$
$$7.26 \le \sigma^2_{\tau\beta} \le 137.81$$

Este resultado es consistente con los resultados de la prueba F exacta para $\sigma^2_{\tau\beta}$, en que hay evidencia sólida de que este componente de la varianza es diferente de cero.

12-7.2 Método de grandes muestras modificado

El método de Satterthwaite de la sección anterior es una forma relativamente simple de encontrar un intervalo de confianza aproximado para un componente de la varianza que puede expresarse como una combinación lineal de cuadrados medios, por ejemplo

$$\hat{\sigma}^2_0 = \sum_{i=1}^{Q} c_i MS_i \tag{12-41}$$

El método de Satterthwaite funciona bien cuando los grados de libertad de cada cuadrado medio MS_i son relativamente grandes, y cuando todas las constantes c_i de la ecuación 12-41 son positivas. Sin embargo, en ocasiones algunas de las c_i son negativas. Graybill y Wang [51] propusieron un procedimiento llamado **método de grandes muestras modificado,** que puede ser una alternativa muy útil del método de Satterthwaite. Si todas las constantes c_i de la ecuación 12-41 son positivas, entonces el intervalo de confianza modificado de $100(1-\alpha)$ por ciento de una muestra grande para σ^2_0 es

$$\hat{\sigma}^2_0 - \sqrt{\sum_{i=1}^{Q} G_i^2 c_i^2 MS_i^2} \le \sigma^2_0 \le \hat{\sigma}^2_0 + \sqrt{\sum_{i=1}^{Q} H_i^2 c_i^2 MS_i^2} \tag{12-42}$$

donde

$$G_i = 1 - \frac{1}{F_{\alpha, f_i, \infty}} \quad \text{y} \quad H_i = \frac{1}{F_{1-\alpha, f_i, \infty}} - 1$$

Observe que una variable aleatoria F con un número infinito de grados de libertad en el denominador es equivalente a una variable aleatoria ji-cuadrada dividida por sus grados de libertad.

Considere ahora el caso más general de la ecuación 12-41, donde las constantes c_i no tienen restricciones sobre el signo. Esto puede escribirse como

$$\hat{\sigma}_0^2 = \sum_{i=1}^{P} c_i MS_i - \sum_{j=P+1}^{Q} c_j MS_j, \qquad c_i,\ c_j \geq 0 \tag{12-43}$$

Ting, *et al* [110], dan un intervalo de confianza inferior aproximado de $100(1-\alpha)$ por ciento para σ_0^2 como

$$L = \hat{\sigma}_0^2 - \sqrt{V_L} \tag{12-44}$$

donde

$$V_L = \sum_{i=1}^{P} G_i^2 c_i^2 MS_i^2 + \sum_{j=P+1}^{Q} H_j^2 c_j^2 MS_j^2 + \sum_{i=1}^{P} \sum_{j=P+1}^{Q} G_{ij}^2 c_i c_j MS_i MS_j$$

$$+ \sum_{i=1}^{P-1} \sum_{t>i}^{P} G_{it}^* c_i c_t MS_i MS_t$$

$$G_i = 1 - \frac{1}{F_{\alpha, f_i, \infty}}$$

$$H_j = \frac{1}{F_{1-\alpha, f_i, \infty}} - 1$$

$$G_{ij} = \frac{(F_{\alpha, f_i, f_j} - 1)^2 - G_i^2 F_{\alpha, f_i, f_j}^2 - H_j^2}{F_{\alpha, f_i, f_j}}$$

$$G_{it}^* = \left[\left(1 - \frac{1}{F_{\alpha, f_i + f_t, \infty}} \right)^2 \frac{(f_i + f_t)^2}{f_i f_t} - \frac{G_i^2 f_i}{f_t} - \frac{G_t^2 f_t}{f_i} \right] (P-1),$$

$$\text{si } P > 1 \quad \text{y} \quad G_{it}^* = 0 \quad \text{si} \quad P = 1$$

Estos resultados también pueden extenderse para incluir intervalos de confianza aproximados para cocientes de componentes de la varianza. Para una relación completa de estos métodos, referirse al excelente libro de Burdick y Graybill [22].

EJEMPLO 12-9 ....

Para ilustrar el método de grandes muestras modificado, considere nuevamente el modelo mixto con tres factores del ejemplo 12-7. Se encontrará un intervalo de confianza inferior aproximado de 95% para $\sigma_{\tau\beta}^2$. Recuerde que la estimación puntual de $\sigma_{\tau\beta}^2$ es

$$\hat{\sigma}_{\tau\beta}^2 = \frac{MS_{AB} - MS_{ABC}}{cn} = \frac{134.91 - 19.26}{(3)(2)} = 19.28$$

Por lo tanto, en la notación de la ecuación 12-43, $c_1 = c_2 = {}^1/_6$, y

$$G_1 = 1 - \frac{1}{F_{0.05,6,\infty}} = 1 - \frac{1}{2.1} = 0.524$$
$$H_2 = \frac{1}{F_{0.95,12,\infty}} - 1 = \frac{1}{0.435} - 1 = 1.30$$
$$G_{12} = \frac{(F_{0.05,6,12} - 1)^2 - (G_1)^2 F^2_{0.05,6,12} - (H_2)^2}{F_{0.05,6,12}}$$
$$= \frac{(3.00-1)^2 - (0.524)^2(3.00)^2 - (1.3)^2}{3.00} = -0.054$$
$$G^*_{1t} = 0$$

Por la ecuación 12-44

$$V_L = G_1^2 c_1^2 MS_{AB}^2 + H_2^2 c_2^2 MS_{ABC}^2 + G_{12} c_1 c_2 MS_{AB} MS_{ABC}$$
$$= (0.524)^2 (1/6)^2 (134.91)^2 + (1.3)^2 (1/6)^2 (19.26)^2$$
$$+ (-0.054)(1/6)(1/6)(134.91)(19.26)$$
$$= 152.36$$

Así, un límite de confianza inferior aproximado de 95% para $\sigma^2_{\tau\beta}$ es

$$L = \hat{\sigma}^2_{\tau\beta} - \sqrt{V_L} = 19.28 - \sqrt{152.36} = 6.94$$

Este resultado es consistente con los resultados de la prueba F exacta para este efecto.

12-7.3 Estimación de máxima verosimilitud de componentes de la varianza

En este capítulo se ha subrayado el método del análisis de varianza para estimar los componentes de la varianza debido a que es relativamente directo y hace uso de cantidades familiares: los cuadrados medios de la tabla del análisis de varianza. Sin embargo, el método tiene ciertas desventajas, incluyendo la molesta tendencia a producir en ocasiones estimaciones negativas. Además, el método del análisis de varianza es en realidad un **método de estimador de momentos,** una técnica que los especialistas en estadística matemática prefieren en general no usar para estimar parámetros, debido a que resulta con frecuencia en estimaciones de parámetros que no tienen buenas propiedades estadísticas. A la técnica de estimación de parámetros preferida se le llama **método de máxima verosimilitud.** La implementación de este método puede ser un tanto complicada, en particular para el modelo de un diseño experimental, pero en cierto sentido el método de máxima verosimilitud selecciona estimaciones de los parámetros que, para un modelo y una distribución del error especificados, maximiza la probabilidad de ocurrencia de los resultados muestrales. Una descripción general muy adecuada del método de máxima verosimilitud aplicado a modelos de diseños experimentales se ofrece en Milliken y Johnson [79].

La revisión completa del método de máxima verosimilitud sale del alcance de este libro, pero la idea general puede ilustrarse con suma facilidad. Soponga que x es una variable aleatoria con una distribución

de probabilidad $f(x; \theta)$, donde θ es un parámetro desconocido. Sea $x_1, x_2, \ldots, x_n$ una muestra aleatoria de n observaciones. Entonces la **función de verosimilitud** de la muestra es

$$L(\theta) = f(x_1;\, \theta) \cdot f(x_2; \theta) \cdot \cdots \cdot f(x_n;\, \theta)$$

Observe que ahora la función de verosimilitud es una función únicamente del parámetro desconocido θ. El estimador de máxima verosimilitud de θ es el valor de θ que maximiza la función de verosimilitud $L(\theta)$.

Para ilustrar cómo se aplica esto en el modelo de un diseño experimental con efectos aleatorios, considere un modelo de dos factores con $a = b = n = 2$. El modelo es

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \varepsilon_{ijk}$$

con $i = 1, 2, j = 1, 2$ y $k = 1, 2$. La varianza de cualquier observación es

$$V(y_{ijk}) = \sigma_y^2 = \sigma_\tau^2 + \sigma_\beta^2 + \sigma_{\tau\beta}^2 + \sigma^2$$

y las covarianzas son

$$\begin{aligned} \text{Cov}(y_{ijk},\, y_{i'j'k'}) &= \sigma_\tau^2 + \sigma_\beta^2 + \sigma_{\tau\beta}^2 && i = i',\ j = j',\ k \neq k' \\ &= \sigma_\tau^2 && i = i',\ j \neq j' \\ &= \sigma_\beta^2 && i \neq i',\ j = j' \\ &= 0 && i \neq i',\ j \neq j' \end{aligned} \tag{12-45}$$

Es conveniente considerar las observaciones como un vector 8×1, es decir,

$$\mathbf{y} = \begin{bmatrix} y_{111} \\ y_{112} \\ y_{211} \\ y_{212} \\ y_{121} \\ y_{122} \\ y_{221} \\ y_{222} \end{bmatrix}$$

y las varianzas y covarianzas pueden expresarse como una **matriz de covarianza** 8×8

$$\mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{bmatrix}$$

donde $\mathbf{\Sigma}_{11}$, $\mathbf{\Sigma}_{22}$, $\mathbf{\Sigma}_{12}$ y $\mathbf{\Sigma}_{21} = \mathbf{\Sigma}'_{12}$ son matrices 4×4 definidas de la siguiente manera:

$$\mathbf{\Sigma}_{11} = \mathbf{\Sigma}_{22} = \begin{bmatrix} \sigma_y^2 & \sigma_\tau^2 + \sigma_\beta^2 + \sigma_{\tau\beta}^2 & \sigma_\tau^2 & \sigma_\tau^2 \\ \sigma_\tau^2 + \sigma_\beta^2 + \sigma_{\tau\beta}^2 & \sigma_y^2 & \sigma_\tau^2 & \sigma_\tau^2 \\ \sigma_\tau^2 & \sigma_\tau^2 & \sigma_y^2 & \sigma_\tau^2 + \sigma_\beta^2 + \sigma_{\tau\beta}^2 \\ \sigma_\tau^2 & \sigma_\tau^2 & \sigma_\tau^2 + \sigma_\beta^2 + \sigma_{\tau\beta}^2 & \sigma_y^2 \end{bmatrix}$$

$$\mathbf{\Sigma}_{12} = \begin{bmatrix} \sigma_\beta^2 & \sigma_\beta^2 & 0 & 0 \\ \sigma_\beta^2 & \sigma_\beta^2 & 0 & 0 \\ 0 & 0 & \sigma_\beta^2 & \sigma_\beta^2 \\ 0 & 0 & \sigma_\beta^2 & \sigma_\beta^2 \end{bmatrix}$$

y $\mathbf{\Sigma}_{21}$ es sólo la transpuesta de $\mathbf{\Sigma}_{12}$. Entonces cada observación sigue una distribución normal con varianza σ_y^2, y si se supone que todas las $N = abn$ observaciones tienen una distribución normal conjunta, entonces la función de verosimilitud del modelo aleatorio queda como

$$L(\mu, \sigma_\tau^2, \sigma_\beta^2, \sigma_{\tau\beta}^2, \sigma^2) = \frac{1}{(2\pi)^{N/2}|\mathbf{\Sigma}|^{1/2}} \exp\left[-\frac{1}{2}(\mathbf{y} - \mathbf{j}_N\mu)'\mathbf{\Sigma}^{-1}(\mathbf{y} - \mathbf{j}_N\mu)\right]$$

donde $\mathbf{j}_N$ es un vector $N \times 1$ compuesto de unos. Las estimaciones de máxima verosimilitud de μ, σ_τ^2, σ_β^2, $\sigma_{\tau\beta}^2$ y σ^2 son los valores de estos parámetros que maximizan la función de verosimilitud. También sería deseable restringir las estimaciones de los componentes de la varianza a valores no negativos. Por lo tanto, en la práctica la función de verosimilitud se maximizaría sujeta a esta restricción.

La estimación de los componentes de la varianza por el método de máxima verosimilitud requiere software de computadora especializado. Algunos paquetes de software de estadística general cuentan con esta capacidad. El sistema SAS calcula estimaciones de máxima verosimilitud de los componentes de la varianza de modelos aleatorios o mixtos con la rutina SAS PROC MIXED. Se ilustrará el uso de la rutina PROC MIXED aplicándola al modelo factorial de dos factores introducido en los ejemplos 12-2 y 12-3.

Considere primero el ejemplo 12-2. Se trata del modelo de un diseño factorial de efectos aleatorios con dos factores. El método del análisis de varianza ha producido una estimación negativa del componente de la varianza de la interacción. Las estimaciones negativas de los componentes de la varianza pueden evitarse en la rutina PROC MIXED especificando el uso del método **de máxima verosimilitud restringida** (o **residual**) (**REML,** por sus siglas en inglés). En esencia, la REML restringe las estimaciones de los componentes de la varianza a valores no negativos.

La rutina PROC MIXED del sistema SAS requiere como entrada la matriz de covarianza de los parámetros del modelo. La estructura de un modelo aleatorio en el que todas las variables aleatorias son mutuamente independientes es

$$\mathbf{G} = \begin{bmatrix} \sigma_\tau^2\mathbf{I} & 0 & 0 \\ 0 & \sigma_\beta^2\mathbf{I} & 0 \\ 0 & 0 & \sigma_{\tau\beta}^2\mathbf{I} \end{bmatrix} \qquad (12\text{-}46)$$

donde las $\mathbf{I}$ son matrices identidad. (La estructura de la covarianza de un modelo puede especificarse en la rutina PROC MIXED con la opción TYPE = *structure* en el enunciado RANDOM.) La estructura de la covarianza del modelo del ejemplo 12-2 se especifica como TYPE = SIM (el valor por omisión de PROC MIXED), que especifica la estructura simple de la covarianza para los parámetros del modelo dados en la ecuación 12-46.

En la tabla 12-17 se presenta la salida de la rutina PROC MIXED de SAS para el experimento del ejemplo 12-2. Se especificó el método de estimación REML de los componentes de la varianza. La salida se ha anotado con números para facilitar la descripción que se presenta a continuación:

1. Estimaciones de los componentes de la varianza y la salida relacionada.
2. Parámetro covarianza. Identifica los parámetros del modelo: σ_τ^2, σ_β^2, $\sigma_{\tau\beta}^2$ y σ^2.
3. Cociente de la varianza estimada del efecto y la varianza estimada del error residual:

$$\hat{\sigma}_i^2 / \hat{\sigma}^2$$

4. Estimaciones de los parámetros. Son las estimaciones REML de los componentes de la varianza $\hat{\sigma}_\tau^2$, $\hat{\sigma}_\beta^2$, $\hat{\sigma}_{\tau\beta}^2$ y $\hat{\sigma}^2$. Observe que la estimación REML de $\hat{\sigma}_{\tau\beta}^2$ es cero.

Tabla 12-17 Salida de PROC MIXED del sistema SAS del análisis del estudio de repetibilidad y reproductibilidad de instrumentos de medición (calibradores) del ejemplo 12-2 utilizando la estimación REML de los componentes de varianza

```
                              The MIXED Procedure

                           Class Level Information
                       Class      Levels  Values
                       PART           20  1 2 3 4 5 6 7 8 9 10 11 12 13
                                          14 15 16 17 18 19 20
                       OPERATOR        3  1 2 3
                       REPLICAT        2  1 2
                                     [1]
                     Covariance Parameter Estimates (REML)
[2]                 [3]          [4]          [5]        [6]    [7]      [8]       [9]
Cov Parm          Ratio       Estimate   Std Errors      Z  Pr > |z| Alpha    Lower    Upper

OPERATOR        0.01203539   0.01062922  0.03286000   0.32   0.7463   0.05  -0.0538   0.0750
PART           11.60743820  10.25126446  3.37376878   3.04   0.0024   0.05   3.6388  16.8637
PART*OPERATOR  -0.00000000  -0.00000000           .      .        .      .        .        .
Residual        1.00000000   0.88316339  0.12616620   7.00   0.0000   0.05   0.6359   1.1304
                                     [10]
                    Asymptotic Covariance Matrix of Estimates

    Cov Parm           OPERATOR          PART  PART*OPERATOR      Residual
    OPERATOR         0.00107978    0.00006632     0.00000000   -0.00039795
    PART             0.00006632   11.38231579     0.00000000   -0.00265287
    PART*OPERATOR    0.00000000    0.00000000     0.00000000   -0.00000000
    Residual        -0.00039795   -0.00265287    -0.00000000    0.01591791
                                     [11]
                       Model Fitting Information for VALUE
                     Description                            Value
                     Observations                        120.0000
                     Variance Estimate                     0.8832
                     Standard Deviation Estimate           0.9398
                     REML Log Likelihood                 -204.696
                     Akaike's Information Criterion      -208.696
                     Schwarz's Bayesian Criterion        -214.254
                     -2 REML Log Likelihood               409.3913
```

Tabla 12-18 Salida de PROC MIXED del sistema SAS del análisis del estudio de repetibilidad y reproductibilidad de instrumentos de medición (calibradores) con el operador como efecto fijo utilizando la estimación REML de los componentes de la varianza

The MIXED Procedure

Covariance Parameter Estimates (REML)

Cov Parm	Ratio	Estimate	Std Error	Z	Pr > |z|	Alpha	Lower	Upper
PART	11.60743876	10.25126472	3.37376895	3.04	0.0024	0.05	3.6388	16.8637
PART*OPERATOR	0.00000000	0.00000000	.	.	.	.	.	.
Residual	1.00000000	0.88316337	0.12616620	7.00	0.0000	0.05	0.6359	1.1304

Asymptotic Covariance Matrix of Estimates

Cov Parm	PART	PART*OPERATOR	Residual
PART	11.38231693	0.00000000	-0.00265287
PART*OPERATOR	0.00000000	0.00000000	-0.00000000
Residual	-0.00265287	0.00000000	-0.01591791

Model Fitting Information for VALUE

Description	Value
Observations	120.0000
Variance Estimate	0.8832
Standard Deviation Estimate	0.9398
REML Log Likelihood	-204.729
Akaike's Information Criterion	-207.729
Schwarz's Bayesian Criterion	-211.872
-2 REML Log Likelihood	409.4572

Tests of Fixed Effects

Source	NDF	DDF	Type III F	Pr > F
OPERATOR	2	38	1.48	0.2401

5. Error estándar de la estimación. Es el error estándar (*se*) para muestras grandes de la estimación del parámetro: $se(\hat{\sigma}_i^2) = \sqrt{V(\hat{\sigma}_i^2)}$.
6. El estadístico Z asociado con la varianza estimada:
$$Z = \hat{\sigma}_i^2 / se(\hat{\sigma}_i^2).$$
7. Valor P del estadístico Z calculado.
8. Nivel alfa usado para calcular el intervalo de confianza.
9. Límites inferior y superior de un intervalo de confianza de la teoría normal para muestras grandes de $100(1 - \alpha)$ por ciento para los componentes de la varianza:
$$L = \hat{\sigma}_i^2 - Z_{\alpha/2} se(\hat{\sigma}_i^2) \qquad U = \hat{\sigma}_i^2 + Z_{\alpha/2} se(\hat{\sigma}_i^2)$$
10. Matriz asintótica de la covarianza de las estimaciones. Es la matriz de la covarianza para muestras grandes de las estimaciones de los componentes de la varianza.
11. Medidas del ajuste del modelo para comparar el ajuste de modelos alternativos.

Observe que los resultados de la rutina PROC MIXED de SAS coinciden muy de cerca con los valores presentados en el ejemplo 12-2 cuando el modelo reducido (sin el término de la interacción) se ajustó a los datos.

En el ejemplo 12-3 se consideró el mismo experimento, pero se supuso que los operadores eran un factor fijo, lo cual llevó a un modelo mixto. La rutina PROC MIXED de SAS puede emplearse para estimar los componentes de la varianza para esta situación. La estructura de la covarianza de las observaciones, suponiendo que todas las variables aleatorias son mutuamente independientes (es decir, el modelo mixto no restringido), es

$$\begin{aligned} \text{Cov}(y_{ijk}, y_{i'j'k'}) &= \sigma_\beta^2 + \sigma_{\tau\beta}^2 + \sigma^2 & i = i',\ j = j',\ k = k' \\ &= \sigma_\beta^2 + \sigma_{\tau\beta}^2 & i = i',\ j = j',\ k \neq k' \\ &= \sigma_\beta^2 & i \neq i',\ j = j' \\ &= 0 & j \neq j' \end{aligned} \tag{12-47}$$

La matriz de la covarianza de los parámetros del modelo es

$$\mathbf{G} = \begin{bmatrix} \sigma_\beta^2 \mathbf{I} & 0 \\ 0 & \sigma_{\tau\beta}^2 \mathbf{I} \end{bmatrix} \tag{12-48}$$

(Esto se especifica en el enunciado TYPE = SIM en la entrada de la rutina PROC MIXED.) En la tabla 12-18 se muestra la salida de la rutina PROC MIXED de SAS para la forma no restringida del modelo mixto del ejemplo 12-3. De nueva cuenta se seleccionó el método REML. La estimación del componente de la varianza para el factor "pieza" es muy similar a la estimación que se obtuvo utilizando el modelo aleatorio. La estimación de la varianza del error residual también es similar. Además, la salida incluye una prueba F para el efecto fijo.

12-8 PROBLEMAS

12-1. Una fábrica textil tiene un gran número de telares. Se supone que cada telar produce la misma cantidad de tela por minuto. Para investigar este supuesto, se eligen cinco telares al azar y se registra su producción en tiempos diferentes. Se obtienen los siguientes datos:

Telar	Producción (lb/min)				
1	14.0	14.1	14.2	14.0	14.1
2	13.9	13.8	13.9	14.0	14.0
3	14.1	14.2	14.1	14.0	13.9
4	13.6	13.8	14.0	13.9	13.7
5	13.8	13.6	13.9	13.8	14.0

a) Explicar por qué este experimento es de efectos aleatorios. ¿Todos los telares tienen la misma producción? Utilizar $\alpha = 0.05$.
b) Estimar la variabilidad entre los telares.
c) Estimar la varianza del error experimental.
d) Encontrar un intervalo de confianza de 95% para $\sigma_\tau^2 / (\sigma_\tau^2 + \sigma^2)$.
e) Analizar los residuales de este experimento. ¿Considera el lector que se satisfacen los supuestos del análisis de varianza?

12-2. Un fabricante sospecha que los lotes de materia prima suministrados por su proveedor difieren de manera significativa en el contenido de calcio. Hay un gran número de lotes actualmente en el almacén. Se seleccionan cinco de ellos para hacer un estudio. Un químico hace cinco determinaciones en cada lote y obtiene los siguientes datos:

Lote 1	Lote 2	Lote 3	Lote 4	Lote 5
23.46	23.59	23.51	23.28	23.29
23.48	23.46	23.64	23.40	23.46
23.56	23.42	23.46	23.37	23.37
23.39	23.49	23.52	23.46	23.32
23.40	23.50	23.49	23.39	23.38

a) ¿Existe una variación significativa en el contenido de calcio de un lote a otro? Utilizar $\alpha = 0.05$.
b) Estimar los componentes de la varianza.
c) Encontrar un intervalo de confianza de 95% para $\sigma_\tau^2 / (\sigma_\tau^2 + \sigma^2)$.
d) Analizar los residuales de este experimento. ¿Se satisfacen los supuestos del análisis de varianza?

12-3. En una fábrica metalúrgica se usan varios hornos para calentar ejemplares de metal. Se supone que todos los hornos operan a la misma temperatura, aunque se sospecha que quizá no sea éste el caso. Se seleccionan al azar tres hornos y se registran sus temperaturas en cargas sucesivas. Los datos recabados son los siguientes:

Horno	Temperatura					
1	491.50	498.30	498.10	493.50	493.60	
2	488.50	484.65	479.90	477.35		
3	490.10	484.80	488.25	473.00	471.85	478.65

a) ¿Existe una variación significativa de la temperatura entre los hornos? Utilizar $\alpha = 0.05$.
b) Estimar los componentes de la varianza de este modelo.
c) Analizar los residuales de este experimento y sacar conclusiones acerca de la adecuación del modelo.

12-4. En un artículo de *Journal of the Electrochemical Society* (vol. 139, no. 2, pp. 524-532) se describe un experimento para investigar la deposición de vapor a baja presión del polisilicio. El experimento se llevó a cabo en el reactor de alta capacidad de Sematech en Austin, Texas. El reactor tiene varias posiciones para las obleas, y se seleccionan al azar cuatro de estas posiciones. La variable de respuesta es la uniformidad del espesor de la película. Se hicieron tres réplicas del experimento y se obtuvieron los siguientes datos:

Posición de la oblea	Uniformidad		
1	2.76	5.67	4.49
2	1.43	1.70	2.19
3	2.34	1.97	1.47
4	0.94	1.36	1.65

a) ¿Hay alguna diferencia en las posiciones de las obleas? Utilizar $\alpha = 0.05$.
b) Estimar la variabilidad debida a las posiciones de las obleas.
c) Estimar el componente del error aleatorio.
d) Analizar los residuales de este experimento y comentar la adecuación del modelo.

12-5. Considere el experimento de la deposición de vapor del problema 12-4.
a) Estimar la variabilidad total de la respuesta uniformidad.
b) ¿Qué parte de la variabilidad total de la respuesta uniformidad se debe a la diferencia entre las posiciones en el reactor?
c) ¿Hasta qué nivel podría reducirse la variabilidad de la respuesta uniformidad si pudiera eliminarse la variabilidad entre una posición y otra en el reactor? ¿Considera el lector que ésta es una reducción significativa?

12-6. En un artículo de *Journal of Quality Technology* (vol. 13, no. 2, pp. 111-114) se describe un experimento para investigar los efectos de cuatro sustancias químicas blanqueadoras sobre la brillantez de la pulpa. Estas cuatro sustancias químicas se seleccionaron al azar de una población grande de agentes blanqueadores potenciales. Los datos son los siguientes:

Sustancia química	Brillantez de la pulpa				
1	77.199	74.466	92.746	76.208	82.876
2	80.522	79.306	81.914	80.346	73.385
3	79.417	78.017	91.596	80.802	80.626
4	78.001	78.358	77.544	77.364	77.386

a) ¿Existe alguna diferencia en los tipos de sustancias químicas? Utilizar $\alpha = 0.05$.
b) Estimar la variabilidad debida al tipo de sustancias químicas.
c) Estimar la variabilidad debida al error aleatorio.
d) Analizar los residuales de este experimento y comentar la adecuación del modelo.

12-7. Considere el modelo de efectos aleatorios, balanceado, en una variable. Desarrollar un procedimiento para encontrar un intervalo de confianza de $100(1 - \alpha)$ por ciento para $\sigma^2/(\sigma_\tau^2 + \sigma^2)$.

12-8. Referirse al problema 12-1.
a) ¿Cuál es la probabilidad de aceptar H_0 si σ_τ^2 es 4 veces la varianza del error σ^2?
b) Si la diferencia entre los telares es lo suficientemente grande para incrementar la desviación estándar de una observación en 20%, quiere detectarse esto con una probabilidad de al menos 0.80. ¿Qué tamaño de la muestra deberá usarse?

12-9. Se llevó a cabo un experimento para investigar la capacidad o aptitud de un sistema de medición. Se seleccionaron diez piezas al azar, y dos operadores escogidos aleatoriamente midieron tres veces cada pieza. Las pruebas se hicieron en orden aleatorio y se obtuvieron los siguientes datos:

Número	Mediciones del operador 1			Mediciones del operador 2		
de pieza	1	2	3	1	2	3
1	50	49	50	50	48	51
2	52	52	51	51	51	51
3	53	50	50	54	52	51
4	49	51	50	48	50	51
5	48	49	48	48	49	48
6	52	50	50	52	50	50
7	51	51	51	51	50	50
8	52	50	49	53	48	50
9	50	51	50	51	48	49
10	47	46	49	46	47	48

a) Analizar los datos de este experimento.
b) Encontrar estimaciones puntuales de los componentes de la varianza utilizando el método del análisis de varianza.

12-10. Considere nuevamente los datos del problema 5-6. Suponga que ambos factores, las máquinas y los operadores, se eligen al azar.
a) Analizar los datos de este experimento.
b) Encontrar estimaciones puntuales de los componentes de la varianza utilizando el método del análisis de varianza.

12-11. Considere nuevamente los datos del problema 5-13. Suponga que ambos factores son aleatorios.
a) Analizar los datos de este experimento.
b) Estimar los componentes de la varianza.

12-12. Suponga que en el problema 5-11 las posiciones en el horno se seleccionaron aleatoriamente, dando como resultado un experimento con un modelo mixto. Analizar de nuevo los datos de este experimento bajo este nuevo supuesto. Estimar los componentes apropiados del modelo.

12-13. Analizar de nuevo el experimento de los sistemas de medición del problema 12-9, suponiendo que los operadores son un factor fijo. Estimar los componentes apropiados del modelo.

12-14. En el problema 5-6, suponga que sólo hay cuatro máquinas de interés, pero los operadores se seleccionaron aleatoriamente.
a) ¿Qué tipo de modelo es apropiado?
b) Efectuar el análisis y estimar los componentes del modelo.

12-15. Mediante la aplicación del operador valor esperado, desarrollar los cuadrados medios esperados del modelo factorial mixto con dos factores. Usar los supuestos del modelo restringido. Verificar los resultados con los cuadrados medios esperados de la tabla 12-11 para constatar que concuerdan.

12-16. Considere el diseño factorial de tres factores del ejemplo 12-6. Proponer los estadísticos de prueba apropiados para todos los efectos principales y las interacciones. Repetir para el caso en que A y B son fijos y C es aleatorio.

12-17. Considere el experimento del ejemplo 12-7. Analizar los datos para el caso en que A, B y C son aleatorios.

12-18. Deducir los cuadrados medios esperados de la tabla 12-14.

12-19. Considere un experimento factorial de cuatro factores donde el factor A tiene a niveles, el factor B tiene b niveles, el factor C tiene c niveles, el factor D tiene d niveles y hay n réplicas. Anotar las sumas de cuadrados, los grados de libertad y los cuadrados medios esperados para los siguientes casos. Suponer el modelo restringido para todos los modelos mixtos. Puede usarse un paquete de computadora como Minitab.
a) A, B, C y D son factores fijos.
b) A, B, C y D son factores aleatorios.
c) A es fijo y B, C y D son aleatorios.
d) A y B son fijos y C y D son aleatorios.
e) A, B y C son fijos y D es aleatorio.
¿Existen pruebas exactas para todos los efectos? Si no es así, proponer estadísticos de prueba para los efectos que no puedan probarse directamente.

12-20. Considere nuevamente los incisos *c*, *d* y *e* del problema 12-19. Obtener los cuadrados medios esperados suponiendo un modelo no restringido. Puede usarse un paquete de computadora como Minitab. Comparar los resultados obtenidos con los del modelo restringido.

12-21. En el problema 5-17, suponga que los tres operadores se seleccionaron al azar. Analizar los datos bajo estas condiciones y sacar conclusiones. Estimar los componentes de la varianza.

12-22. Considere el modelo factorial de tres factores

$$y_{ijk} = \mu + \tau_i + \beta_j + \gamma_k + (\tau\beta)_{ij} + (\beta\gamma)_{jk} + \varepsilon_{ijk} \qquad \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, b \\ k = 1, 2, \ldots, c \end{cases}$$

Suponiendo que los tres factores son aleatorios, desarrollar la tabla del análisis de varianza, incluyendo los cuadrados medios esperados. Proponer los estadísticos de prueba apropiados para todos los efectos.

12-23. El modelo factorial de tres factores para una sola réplica es

$$y_{ijk} = \mu + \tau_i + \beta_j + \gamma_k + (\tau\beta)_{ij} + (\beta\gamma)_{jk} + (\tau\gamma)_{ik} + (\tau\beta\gamma)_{ijk} + \varepsilon_{ijk}$$

Si todos los factores son aleatorios, ¿puede probarse alguno de los efectos? Si las interacciones de tres factores y $(\tau\beta)_{ij}$ no existen, ¿es posible probar todos los demás efectos?

12-24. En el problema 5-6, suponga que tanto las máquinas como los operadores se escogieron al azar. Determinar la potencia de la prueba para detectar un efecto de la máquina tal que $\sigma_\beta^2 = \sigma^2$, donde σ_β^2 es el componente de la varianza del factor máquina. ¿Son suficientes dos réplicas?

12-25. En el análisis de varianza del modelo mixto de dos factores, demostrar que $\text{Cov}[(\tau\beta)_{ij}, (\tau\beta)_{i'j}] = -(1/a)\,\sigma_{\tau\beta}^2$ para $i \neq i'$.

12-26. Demostrar que el método del análisis de varianza siempre produce estimaciones puntuales insesgadas de los componentes de la varianza en cualquier modelo aleatorio o mixto.

12-27. Invocando los supuestos de normalidad usuales, encontrar una expresión para la probabilidad de obtener una estimación negativa de un componente de la varianza por el método del análisis de varianza. Utilizando este resultado, escribir un enunciado para la probabilidad de que $\hat{\sigma}_\tau^2 < 0$ en un análisis de varianza de un factor. Comentar la utilidad de este enunciado de probabilidad.

12-28. Analizar los datos del problema 12-9, suponiendo que los operadores son fijos y utilizando tanto la forma no restringida como la restringida de los modelos mixtos. Comparar los resultados que se obtienen con los dos modelos.

12-29. Considere el modelo mixto de dos factores. Demostrar que el error estándar de la media del factor fijo (por ejemplo, A) es $[MS_{AB}/bn]^{1/2}$.

12-30. Considere los componentes de la varianza del modelo aleatorio del problema 12-9.

a) Encontrar un intervalo de confianza exacto del 95% para σ^2.

b) Encontrar intervalos de confianza aproximados del 95% para los otros componentes de la varianza utilizando el método de Satterthwaite.

12-31. Usar el experimento descrito en el problema 5-6 y suponer que ambos factores son aleatorios. Encontrar un intervalo de confianza exacto del 95% para σ^2. Construir intervalos de confianza aproximados del 95% para los otros componentes de la varianza utilizando el método de Satterthwaite.

12-32. Considere el experimento de tres factores del problema 5-17 y suponga que los operadores se seleccionaron al azar. Encontrar un intervalo de confianza aproximado del 95% para el componente de la varianza del operador.

12-33. Resolver de nuevo el problema 12-30 utilizando el método de grandes muestras modificado que se describe en la sección 12-7.2. Comparar las dos series de intervalos de confianza obtenidas y comentarlas.

12-34. Resolver de nuevo el problema 12-32 utilizando el método de grandes muestras modificado que se describe en la sección 12-7.2. Comparar este intervalo de confianza con el que se obtuvo anteriormente y comentarlo.

13 Diseños anidados y de parcelas subdivididas

En este capítulo se introducen dos importantes tipos de diseños experimentales: el **diseño anidado** y el **diseño de parcelas subdivididas.** Estos dos diseños encuentran una aplicación razonablemente generalizada en el uso industrial de los experimentos diseñados. Con frecuencia incluyen también uno o más **factores aleatorios,** por lo que algunos de los conceptos introducidos en el capítulo 12 tendrán cabida aquí.

13-1 DISEÑO ANIDADO DE DOS ETAPAS

En algunos experimentos con factores múltiples, los niveles de uno de los factores (por ejemplo el factor B) son similares pero no idénticos a los diferentes niveles de otro factor (por ejemplo A). A un arreglo como éste se le llama **diseño anidado** o **jerárquico,** con los niveles del factor B anidados bajo los niveles del factor A. Por ejemplo, considere una compañía que compra su materia prima a tres proveedores diferentes. La compañía quiere determinar si la pureza de la materia prima de cada proveedor es la misma. Hay cuatro lotes de materia prima disponibles de cada proveedor, y se harán tres determinaciones de la pureza en cada lote. La situación se describe en la figura 13-1.

Se trata de un **diseño anidado de dos etapas,** con los lotes anidados bajo los proveedores. A primera vista se podría preguntar por qué no es un experimento factorial. Si fuera un experimento factorial, entonces el lote 1 se referiría siempre al mismo lote, el lote 2 se referiría siempre al mismo lote, etcétera. Evidentemente, no es éste el caso, ya que los lotes de cada proveedor son **únicos** para ese proveedor particular. Es decir, el lote 1 del proveedor 1 no tiene relación con el lote 1 de cualquier otro proveedor, el lote 2 del proveedor 1 no tiene relación con el lote 2 de cualquier otro proveedor, etcétera. Para subrayar el hecho de que los lotes de cada proveedor son diferentes, se pueden numerar como 1, 2, 3 y 4 para el proveedor 1; 5, 6, 7 y 8 para el proveedor 2; y 9, 10, 11 y 12 para el proveedor 3, como se muestra en la figura 13-2.

En ocasiones quizá no se sepa si un factor está cruzado en un arreglo factorial o anidado. Si los niveles del factor pueden numerarse arbitrariamente como en la figura 13-2, entonces el factor está anidado.

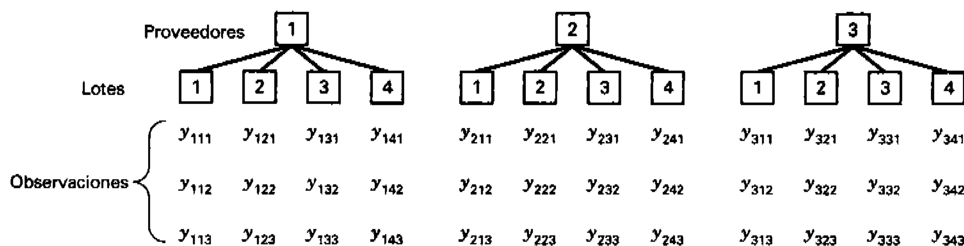


Figura 13-1 Diseño anidado de dos etapas.

13-1.1 Análisis estadístico

El modelo estadístico lineal para el diseño anidado de dos etapas es

$$y_{ijk} = \mu + \tau_i + \beta_{j(i)} + \varepsilon_{(ij)k} \qquad \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, b \\ k = 1, 2, \ldots, n \end{cases} \tag{13-1}$$

Es decir, hay a niveles del factor A, b niveles del factor B anidados bajo cada nivel de A, y n réplicas. El subíndice $j(i)$ indica que el nivel j-ésimo del factor B está anidado bajo el nivel i-ésimo del factor A. Resulta conveniente considerar que las réplicas están anidadas dentro de la combinación de los niveles de A y B; por lo tanto, se usa el subíndice $(ij)k$ para el término del error. Se trata de un **diseño anidado balanceado,** ya que hay el mismo número de niveles de B con cada nivel de A y el mismo número de réplicas. Puesto que no todos los niveles del factor B aparecen dentro de todos los niveles del factor A, no puede haber interacción entre A y B.

La suma de cuadrados total corregida puede escribirse como

$$\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} (y_{ijk} - \bar{y}_{...})^2 = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} [(\bar{y}_{i..} - \bar{y}_{...}) + (\bar{y}_{ij.} - \bar{y}_{i..}) + (y_{ijk} - \bar{y}_{ij.})]^2 \tag{13-2}$$

Al desarrollar el miembro derecho de la ecuación 13-2, se obtiene

$$\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} (y_{ijk} - \bar{y}_{...})^2 = bn\sum_{i=1}^{a} (\bar{y}_{i..} - \bar{y}_{...})^2 + n\sum_{i=1}^{a}\sum_{j=1}^{b} (\bar{y}_{ij.} - \bar{y}_{i..})^2 + \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} (y_{ijk} - \bar{y}_{ij.})^2 \tag{13-3}$$

ya que los tres términos con productos cruzados son cero. La ecuación 13-3 indica que puede hacerse la partición de la suma de cuadrados total en una suma de cuadrados debida al factor A, una suma de cuadrados debida al factor B bajo los niveles de A, y una suma de cuadrados debida al error. Simbólicamente, la ecuación 13-3 puede escribirse como

$$SS_T = SS_A + SS_{B(A)} + SS_E \tag{13-4}$$

Hay $abn - 1$ grados de libertad para SS_T, $a - 1$ grados de libertad para SS_A, $a(b-1)$ grados de libertad para $SS_{B(A)}$ y $ab(n-1)$ grados de libertad para el error. Observe que $abn - 1 = (a-1) + a(b-1) + ab(n-1)$. Si los errores son NID$(0, \sigma^2)$, cada una de las sumas de cuadrados del miembro derecho de la ecuación 13-4

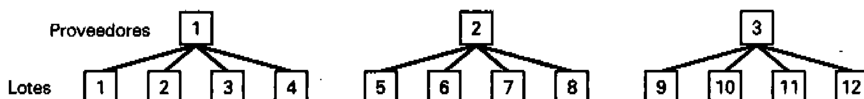


Figura 13-2 Disposición alternativa del diseño anidado de dos etapas.

Tabla 13-1 Cuadrados medios esperados en el diseño anidado de dos etapas

$E(MS)$	A fijo B fijo	A fijo B aleatorio	A aleatorio B aleatorio
$E(MS_A)$	$\sigma^2 + \frac{bn\sum \tau_i^2}{a-1}$	$\sigma^2 + n\sigma_\beta^2 + \frac{bn\sum \tau_i^2}{a-1}$	$\sigma^2 + n\sigma_\beta^2 + bn\sigma_\tau^2$
$E(MS_{B(A)})$	$\sigma^2 + \frac{n\sum\sum \beta_{j(i)}^2}{a(b-1)}$	$\sigma^2 + n\sigma_\beta^2$	$\sigma^2 + n\sigma_\beta^2$
$E(MS_E)$	σ^2	σ^2	σ^2

puede dividirse por sus grados de libertad para obtener cuadrados medios con una distribución independiente tales que el cociente de dos cuadrados medios cualesquiera se distribuye como F.

Los estadísticos apropiados para probar los efectos de los factores A y B dependen de si A y B son **fijos** o **aleatorios.** Si los factores A y B son fijos, se supone que $\Sigma_{i=1}^{a} \tau_i = 0$ y $\Sigma_{j=1}^{b} \beta_{j(i)} = 0$ $(i = 1, 2, ..., a)$. Es decir, la suma de los efectos del tratamiento A es cero, y la suma de los efectos del tratamiento B es cero dentro de cada nivel de A. De manera alternativa, si A y B son aleatorios, se supone que τ_i es NID$(0, \sigma_\tau^2)$ y que $\beta_{j(i)}$ es NID$(0, \sigma_\beta^2)$. También es frecuente encontrar modelos mixtos con A fijo y B aleatorio. Los cuadrados medios esperados pueden determinarse aplicando directamente las reglas del capítulo 12. Para el modelo mixto, estos cuadrados medios esperados suponen la forma **restringida** del modelo del capítulo 12. En la tabla 13-1 se muestran los cuadrados medios esperados para estas situaciones.

La tabla 13-1 indica que si los niveles de A y B son fijos, H_0:$\tau_i = 0$ se prueba con MS_A/MS_E y H_0:$\beta_{j(i)} = 0$ se prueba con $MS_{B(A)}/MS_E$. Si A es un factor fijo y B es aleatorio, entonces H_0:$\tau_i = 0$ se prueba con $MS_A/MS_{B(A)}$ y H_0:$\sigma_\beta^2 = 0$ se prueba con $MS_{B(A)}/MS_E$. Por último, si tanto A como B son factores aleatorios, H_0:$\sigma_\tau^2 = 0$ se prueba con $MS_A/MS_{B(A)}$ y H_0:$\sigma_\beta^2 = 0$ con $MS_{B(A)}/MS_E$. El procedimiento de prueba se resume en la tabla del análisis de varianza, como se muestra en la tabla 13-2. Las fórmulas para calcular las sumas de cuadrados se obtienen desarrollando las cantidades de la ecuación 13-3 y simplificando. Éstas son

$$SS_A = \frac{1}{bn}\sum_{i=1}^{a} y_{i..}^2 - \frac{y_{...}^2}{abn} \tag{13-5}$$

$$SS_{B(A)} = \frac{1}{n}\sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij.}^2 - \frac{1}{bn}\sum_{i=1}^{a} y_{i..}^2 \tag{13-6}$$

$$SS_E = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} y_{ijk}^2 - \frac{1}{n}\sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij.}^2 \tag{13-7}$$

$$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} y_{ijk}^2 - \frac{y_{...}^2}{abn} \tag{13-8}$$

Tabla 13-2 Tabla del análisis de varianza para el diseño anidado de dos etapas

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio
A	$bn\sum (\bar{y}_{i..} - \bar{y}_{...})^2$	$a-1$	MS_A
B dentro de A	$n\sum\sum (\bar{y}_{ij.} - \bar{y}_{i..})^2$	$a(b-1)$	$MS_{B(A)}$
Error	$\sum\sum\sum (y_{ijk} - \bar{y}_{ij.})^2$	$ab(n-1)$	MS_E
Total	$\sum\sum\sum (y_{ijk} - \bar{y}_{...})^2$	$abn-1$	

Se observa que la ecuación 13-6 para $SS_{B(A)}$ puede escribirse como

$$SS_{B(A)} = \sum_{i=1}^{a} \left[\frac{1}{n} \sum_{j=1}^{b} y_{ij.}^2 - \frac{y_{i..}^2}{bn} \right]$$

Esto expresa la idea de que $SS_{B(A)}$ es la suma de cuadrados entre los niveles de B para cada nivel de A, sumados en todos los niveles de A.

EJEMPLO 13-1 ....

Considere una compañía que compra materia prima en lotes de tres proveedores diferentes. La pureza de esta materia prima varía considerablemente, lo cual ocasiona problemas en la manufactura del producto terminado. Quiere determinarse si la variabilidad de la pureza es atribuible a las diferencias entre los proveedores. Se seleccionan al azar cuatro lotes de materia prima de cada proveedor, y se hacen tres determinaciones de la pureza en cada lote. Se trata, desde luego, de un diseño anidado de dos etapas. Los datos, después de codificarlos restando 93, se muestran en la tabla 13-3. Las sumas de cuadrados se calculan de la siguiente manera:

$$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} y_{ijk}^2 - \frac{y_{...}^2}{abn}$$

$$= 153.00 - \frac{(13)^2}{36} = 148.31$$

$$SS_A = \frac{1}{bn}\sum_{i=1}^{a} y_{i..}^2 - \frac{y_{...}^2}{abn}$$

$$= \frac{1}{(4)(3)}[(-5)^2 + (4)^2 + (14)^2] - \frac{(13)^2}{36}$$

$$= 19.75 - 4.69 = 15.06$$

$$SS_{B(A)} = \frac{1}{n}\sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij.}^2 - \frac{1}{bn}\sum_{i=1}^{a} y_{i..}^2$$

$$= \frac{1}{3}[(0)^2 + (-9)^2 + (-1)^2 + \cdots + (2)^2 + (6)^2] - 19.75$$

$$= 89.67 - 19.75 = 69.92$$

y

$$SS_E = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} y_{ijk}^2 - \frac{1}{n}\sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij.}^2$$

$$= 153.00 - 89.67 = 63.33$$

Tabla 13-3 Datos codificados de la pureza del ejemplo 13-1 (Codificación: y_{ijk} = pureza – 93)

		Proveedor 1				Proveedor 2				Proveedor 3			
	Lotes	1	2	3	4	1	2	3	4	1	2	3	4
		1	–2	–2	1	1	0	–1	0	2	–2	1	3
		–1	–3	0	4	–2	4	0	3	4	0	–1	2
		0	–4	1	0	–3	2	–2	2	0	2	2	1
Totales de los lotes	$y_{ij.}$	0	–9	–1	5	–4	6	–3	5	6	0	2	6
Totales de los proveedores	$y_{i..}$	–5				4				14			

Tabla 13-4 Análisis de varianza de los datos del ejemplo 13-1

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	Cuadrado medio esperado	F_0	Valor P
Proveedores	15.06	2	7.53	$\sigma^2 + 3\sigma_\beta^2 + 6\sum \tau_i^2$	0.97	0.42
Lotes (dentro de los proveedores)	69.92	9	7.77	$\sigma^2 + 3\sigma_\beta^2$	2.94	0.02
Error	63.33	24	2.64	σ^2		
Total	148.31	35				

En la tabla 13-4 se resume el análisis de varianza. Los proveedores son fijos y los lotes aleatorios, por lo que los cuadrados medios esperados se obtienen de la columna de en medio de la tabla 13-1 y se repiten por conveniencia en la tabla 13-4. Por el examen de los valores P, se concluiría que no hay ningún efecto significativo sobre la pureza debido a los proveedores, pero la pureza de los lotes de materia prima del mismo proveedor difieren significativamente.

Las implicaciones prácticas de este experimento y del análisis son muy importantes. El objetivo del experimentador es encontrar la fuente de la variabilidad en la pureza de la materia prima. Si ésta es resultado de las diferencias entre los proveedores, el problema puede resolverse seleccionando al "mejor" proveedor. Sin embargo, esa solución no es aplicable aquí porque la principal fuente de variabilidad es la variación de la pureza de un lote a otro *dentro* de los proveedores. Por lo tanto, el problema debe atacarse trabajando con los proveedores para reducir su variabilidad de un lote a otro. Esto puede implicar modificaciones en los procesos de producción de los proveedores o en su sistema interno de control de calidad.

Observe lo que habría pasado si se hubiera hecho un análisis incorrecto de este diseño como un experimento factorial de dos factores. Si se considera que los lotes están cruzados con los proveedores, se obtienen los totales de los lotes de 2, –3, –2 y 16, donde cada celda lote × proveedores contiene tres réplicas. Por lo tanto, puede calcularse una suma de cuadrados debida a los lotes y una suma de cuadrados de interacción. El análisis de varianza factorial completo se muestra en la tabla 13-5, suponiendo un modelo mixto.

Este análisis indica que los lotes difieren significativamente y que hay una interacción significativa entre los lotes y los proveedores. Sin embargo, es difícil ofrecer una interpretación práctica de la interacción lotes × proveedores. Por ejemplo, ¿esta interacción significativa quiere decir que el efecto del proveedor no es constante de un lote a otro? Además, la interacción significativa aunada al efecto no significativo del proveedor podría llevar al analista a concluir que los proveedores en realidad difieren, pero su efecto está enmascarado por la interacción significativa.

Tabla 13-5 Análisis de varianza incorrecto del diseño anidado de dos etapas del ejemplo 13-1 como un diseño factorial (proveedores fijos, lotes aleatorios)

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0	Valor P
Proveedores (S)	15.06	2	7.53	1.02	0.42
Lotes (B)	25.64	3	8.55	3.24	0.04
Interacción $S \times B$	44.28	6	7.38	2.80	0.03
Error	63.33	24	2.64		
Total	148.31	35			

Cálculos

Algunos paquetes de software de estadística realizarán el análisis de un diseño anidado. En la tabla 13-6 se presenta la salida del procedimiento Balanced ANOVA (análisis de varianza balanceado) de Minitab (utilizando el modelo restringido). Los resultados numéricos concuerdan con los cálculos manuales reportados en la tabla 13-4. Minitab también reporta los cuadrados medios esperados en la parte inferior de la tabla 13-6. Recuerde que el símbolo $Q[1]$ es un término cuadrático que representa el efecto fijo de los proveedores, por lo que en la notación que se usa aquí,

$$Q[1] = \frac{\sum_{i=1}^{a} \tau_i^2}{a-1}$$

Por lo tanto, el término del efecto fijo en el cuadrado medio esperado de Minitab para los proveedores $12Q[1] = 12\Sigma_{i=1}^{3} \tau_i^2 / (3-1) = 6\Sigma_{i=1}^{3} \tau_i^2$, resultado que concuerda con el que se presenta en el algoritmo tabular de la tabla 13-4.

En ocasiones no se cuenta con un programa de computadora especializado para analizar diseños anidados. Sin embargo, observe, al comparar las tablas 13-4 y 13-5, que

$$SS_B + SS_{S\times B} = 25.64 + 44.28 = 69.92 \equiv SS_{B(S)}$$

Es decir, la suma de cuadrados de los lotes dentro de los proveedores se compone de la suma de cuadrados de los lotes más la suma de cuadrados de la interacción lotes × proveedores. Los grados de libertad poseen una propiedad similar; es decir,

$$\frac{\text{Lotes}}{3} + \frac{\text{Lotes} \times \text{Proveedores}}{6} = \frac{\text{Lotes dentro de los proveedores}}{9}$$

Por lo tanto, un programa de computadora para analizar diseños factoriales podría usarse también para analizar diseños anidados agrupando el "efecto principal" del factor anidado y las interacciones de ese factor con el factor bajo el que está anidado.

Tabla 13-6 Salida de Minitab (Balanced ANOVA) [análisis de varianza balanceado] para el ejemplo 13-1

Análisis de varianza (diseños balanceados)

```
Factor              Type  Levels Values
Supplier           fixed       3      1      2      3
Batch(Supplier)   random       4      1      2      3      4

Analysis of Variance for Purity

Source              DF         SS         MS        F       P
Supplier             2     15.056      7.528     0.97   0.416
Batch(Supplier)      9     69.917      7.769     2.94   0.017
Error               24     63.333      2.639
Total               35    148.306

Source              Variance Error Expected Mean Square for Each Term
                    component term (using restricted model)
 1 Supplier                      2   (3) + 3(2) + 12Q[1]
 2 Batch(Supplier)      1.710    3   (3) + 3(2)
 3 Error                2.639        (3)
```

13-1.2 Verificación del diagnóstico

La herramienta principal para verificar el diagnóstico es el **análisis residual.** Para el diseño anidado de dos etapas, los residuales son

$$e_{ijk} = y_{ijk} - \hat{y}_{ijk}$$

El valor ajustado es

$$\hat{y}_{ijk} = \hat{\mu} + \hat{\tau}_i + \hat{\beta}_{j(i)}$$

y si se establecen las restricciones usuales sobre los parámetros del modelo ($\Sigma_i\ \hat{\tau}_i = 0$ y $\Sigma_j\ \hat{\beta}_{j(i)} = 0, i = 1, 2, \ldots, a$), entonces $\hat{\mu} = \bar{y}_{...}, \hat{\tau}_i = \bar{y}_{i..} - \bar{y}_{...}$, y $\hat{\beta}_{j(i)} = \bar{y}_{ij.} - \bar{y}_{i..}$. Por consiguiente, el valor ajustado es

$$\begin{aligned}\hat{y}_{ijk} &= \bar{y}_{...} + (\bar{y}_{i..} - \bar{y}_{...}) + (\bar{y}_{ij.} - \bar{y}_{i..}) \\ &= \bar{y}_{ij.}\end{aligned}$$

Por lo tanto, los residuales del diseño anidado de dos etapas son

$$e_{ijk} = y_{ijk} - \bar{y}_{ij.} \tag{13-9}$$

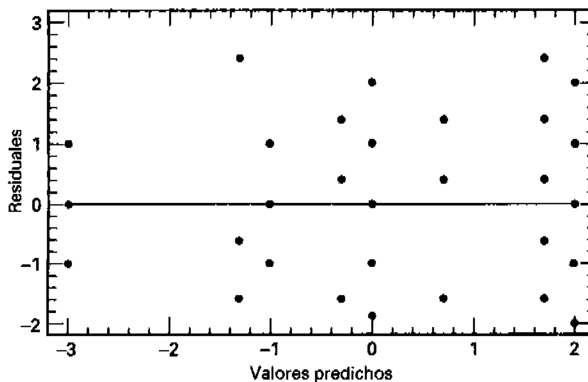
donde $\bar{y}_{ij.}$ son los promedios de los lotes individuales.

Las observaciones, los valores ajustados y los residuales para los datos de la pureza del ejemplo 13-1 son:

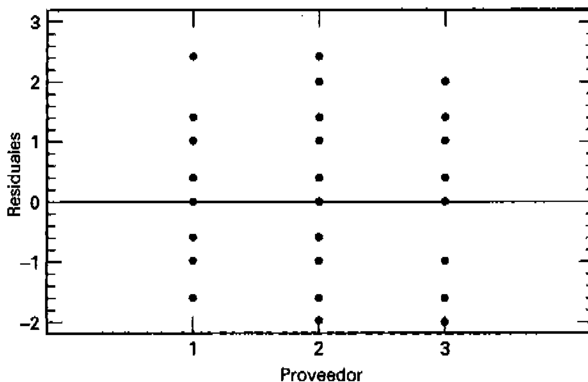
Valor observado y_{ijk}	Valor ajustado $\hat{y}_{ijk} = \bar{y}_{ij.}$	$e_{ijk} = y_{ijk} - \bar{y}_{ij.}$
1	0.00	1.00
−1	0.00	−1.00
0	0.00	0.00
−2	−3.00	1.00
−3	−3.00	0.00
−4	−3.00	−1.00
−2	−0.33	−1.67
0	−0.33	0.33
1	−0.33	1.33
1	1.67	−0.67
4	1.67	2.33
0	1.67	−1.67
1	−1.33	2.33
−2	−1.33	−0.67
−3	−1.33	−1.67
0	2.00	−2.00
4	2.00	2.00
2	2.00	0.00
−1	−1.00	0.00
0	−1.00	1.00
−2	−1.00	−1.00
0	1.67	−1.67
3	1.67	1.33
2	1.67	0.33
2	2.00	0.00

Valor observado y_{ijk}	Valor ajustado $\hat{y}_{ijk} = \bar{y}_{ij.}$	$e_{ijk} = y_{ijk} - \bar{y}_{ij.}$
4	2.00	2.00
0	2.00	–2.00
–2	0.00	–2.00
0	0.00	0.00
2	0.00	2.00
1	0.67	0.33
–1	0.67	–1.67
2	0.67	1.33
3	2.00	1.00
2	2.00	0.00
1	2.00	–1.00

Pueden realizarse ahora las verificaciones de diagnóstico usuales; incluyendo las gráficas de probabilidad normal, la verificación de puntos atípicos y la graficación de los residuales contra los valores ajustados. Como una ilustración, en la figura 13-3 se grafican los residuales contra los valores ajustados y contra los niveles del factor proveedor.



a) Gráfica de los residuales contra los valores predichos



b) Gráfica de los residuales contra el proveedor

Figura 13-3 Gráficas de los residuales del ejemplo 13-1.

En la situación de un problema como el que se describe en el ejemplo 13-1, las gráficas de los residuales son particularmente útiles debido a la información de diagnóstico adicional que contienen. Por ejemplo, el análisis de varianza ha indicado que la pureza media de los tres proveedores no difiere pero que hay una variabilidad estadísticamente significativa de un lote a otro (es decir, $\sigma_\beta^2 > 0$). Pero, ¿la variabilidad dentro de los lotes es la misma para todos los proveedores? Se ha supuesto de hecho que éste es el caso, y si no es cierto desde luego que nos gustaría saberlo, ya que tiene un impacto práctico considerable sobre la interpretación de los resultados del experimento. La gráfica de los residuales contra los proveedores de la figura 13-3*b* es una manera simple pero eficaz de verificar este supuesto. Puesto que la dispersión de los residuales es aproximadamente la misma para los tres proveedores, se concluiría que la variabilidad en la pureza de un lote a otro es aproximadamente la misma para los tres proveedores.

13-1.3 Componentes de la varianza

Para el caso de efectos aleatorios, el método del análisis de varianza puede usarse para estimar los componentes de la varianza σ^2, σ_β^2 y σ_τ^2. Por los cuadrados medios esperados de la última columna de la tabla 13-1, se obtiene

$$\hat{\sigma}^2 = MS_E \qquad (13\text{-}10)$$

$$\hat{\sigma}_\beta^2 = \frac{MS_{B(A)} - MS_E}{n} \qquad (13\text{-}11)$$

y

$$\hat{\sigma}_\tau^2 = \frac{MS_A - MS_{B(A)}}{bn} \qquad (13\text{-}12)$$

En muchas aplicaciones de diseños anidados interviene un **modelo mixto,** con el efecto principal (A) fijo y el factor anidado (B) aleatorio. Éste es el caso para el problema descrito en el ejemplo 13-1; los proveedores (factor A) son fijos, y los lotes de materia prima (factor B) son aleatorios. Los efectos de los proveedores pueden estimarse con

$$\hat{\tau}_1 = \bar{y}_{1..} - \bar{y}_{...} = \frac{-5}{12} - \frac{13}{36} = \frac{-28}{36}$$

$$\hat{\tau}_2 = \bar{y}_{2..} - \bar{y}_{...} = \frac{4}{12} - \frac{13}{36} = \frac{-1}{36}$$

$$\hat{\tau}_3 = \bar{y}_{3..} - \bar{y}_{...} = \frac{14}{12} - \frac{13}{36} = \frac{29}{36}$$

Para estimar los componentes de la varianza σ^2 y σ_β^2, se elimina la línea de la tabla del análisis de varianza relativo a los proveedores y se aplica el método de estimación del análisis de varianza a las dos líneas siguientes. Se obtiene así

$$\hat{\sigma}^2 = MS_E = 2.64$$

y

$$\hat{\sigma}_\beta^2 = \frac{MS_{B(A)} - MS_E}{n} = \frac{7.77 - 2.64}{3} = 1.71$$

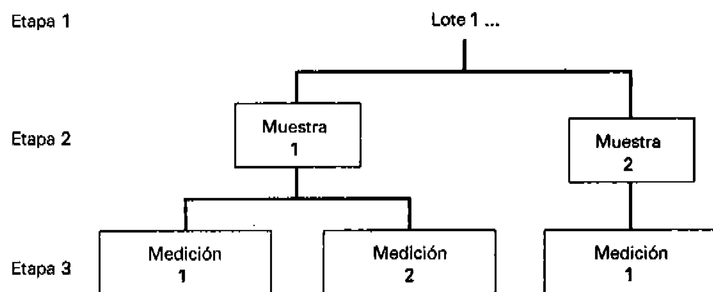


Figura 13-4 Diseño anidado escalonado de dos etapas.

Estos resultados se muestran también en la parte inferior de la salida de Minitab de la tabla 13-6. Por el análisis del ejemplo 13-1, se sabe que τ_i no difiere significativamente de cero, mientras que el componente de la varianza σ_β^2 es mayor que cero.

13-1.4 Diseños anidados por etapas

Un problema potencial en la aplicación de los diseños anidados es que en ocasiones para obtener un número razonable de grados de libertad en el nivel más alto, puede terminarse con muchos grados de libertad (quizá demasiados) en las etapas inferiores. Para ilustrar, suponga que se están investigando las diferencias potenciales en el análisis químico entre diferentes lotes de material. Se planea tomar cinco muestras por lote, y cada muestra se medirá dos veces. Si quiere estimarse un componente de la varianza para los lotes, entonces 10 lotes no sería una elección irrazonable. Esto resulta en 9 grados de libertad para los lotes, 40 grados de libertad para las muestras y 50 grados de libertad para las mediciones.

Una manera de evitar esta situación es usar un tipo particular de diseño anidado no balanceado llamado **diseño anidado por etapas.** En la figura 13-4 se muestra un ejemplo de un diseño anidado escalonado. Observe que sólo se toman dos muestras de cada lote; una de ellas se mide dos veces, mientras que la otra una sola vez. Si hay a lotes, entonces habrá $a - 1$ grados de libertad para los lotes (o, en general, la etapa superior), y todas las etapas inferiores tendrán exactamente a grados de libertad. Para más información sobre el uso y el análisis de estos diseños, ver Bainbridge [5], Smith y Beverly [104] y Nelson [88a, b, c], así como el material suplementario del texto de este capítulo.

13-2 DISEÑO ANIDADO GENERAL DE m ETAPAS

Los resultados de la sección 13-1 pueden extenderse fácilmente al caso de m factores completamente anidados. A este diseño se le llamaría **diseño anidado de m etapas.** Como un ejemplo, suponga que una fundición quiere investigar la dureza de dos formulaciones diferentes de una aleación de metal. Se preparan tres hornadas de cada formulación de la aleación, se seleccionan dos lingotes al azar de cada hornada para probarlos, y se hacen dos mediciones de la dureza en cada lingote. La situación se ilustra en la figura 13-5.

En este experimento, las hornadas están anidadas bajo los niveles del factor formulación de la aleación, y los lingotes están anidados bajo los niveles del factor hornada. Por lo tanto, se trata de un diseño anidado de tres etapas con dos réplicas.

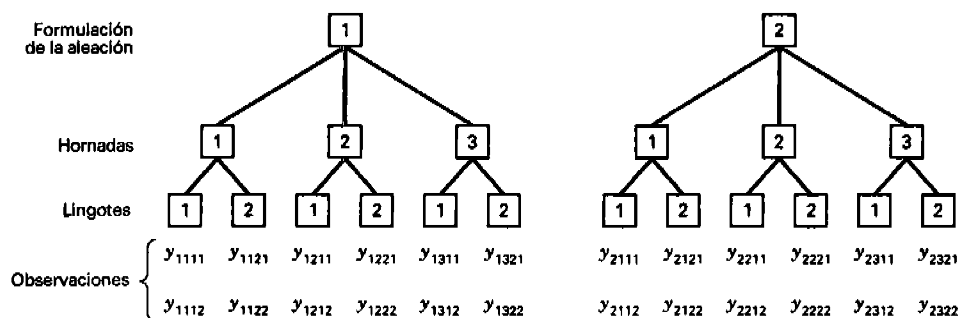


Figura 13-5 Diseño anidado de tres etapas.

El **modelo** para el diseño anidado general de tres etapas es

$$y_{ijkl} = \mu + \tau_i + \beta_{j(i)} + \gamma_{k(ij)} + \varepsilon_{(ijk)l} \qquad \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, b \\ k = 1, 2, \ldots, c \\ l = 1, 2, \ldots, n \end{cases} \tag{13-13}$$

Para el ejemplo tratado aquí, τ_i es el efecto de la formulación de la aleación i-ésima, $\beta_{j(i)}$ es el efecto de la hornada j-ésima dentro de la aleación i-ésima, $\gamma_{k(ij)}$ es el efecto del lingote k-ésimo dentro de la hornada

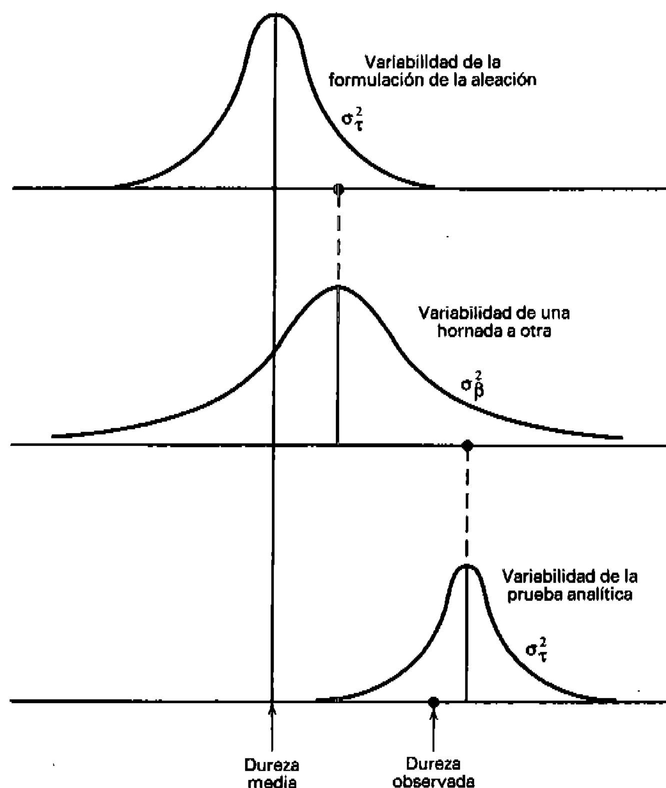


Figura 13-6 Fuentes de variación en el ejemplo del diseño anidado de tres etapas.

Tabla 13-7 Análisis de varianza del diseño anidado de tres etapas

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio
A	$bcn\sum_i (\bar{y}_{i...} - \bar{y}_{....})^2$	$a-1$	MS_A
B (dentro de A)	$cn\sum_i\sum_j (\bar{y}_{ij..} - \bar{y}_{....})^2$	$a(b-1)$	$MS_{B(A)}$
C (dentro de B)	$n\sum_i\sum_j\sum_k (\bar{y}_{ijk.} - \bar{y}_{....})^2$	$ab(c-1)$	$MS_{C(B)}$
Error	$\sum_i\sum_j\sum_k\sum_l (y_{ijkl} - \bar{y}_{ijk.})^2$	$abc(n-1)$	MS_E
Total	$\sum_i\sum_j\sum_k\sum_l (y_{ijkl} - \bar{y}_{....})^2$	$abcn-1$	

j-ésima y la aleación i-ésima, y $\varepsilon_{(ijk)l}$ es el término del error NID$(0, \sigma^2)$ usual. La extensión de este modelo a m factores es directa.

Observe que en el ejemplo anterior la variabilidad global de la dureza constó de tres componentes: uno que resultó de las formulaciones de las aleaciones, otro que se generó de las hornadas y uno más que salió del error de la prueba analítica. Estos componentes de la variabilidad en la dureza global se ilustran en la figura 13-6.

Este ejemplo demuestra la manera en que se usa frecuentemente el diseño anidado en el análisis de procesos para identificar las principales fuentes de variabilidad en la salida. Por ejemplo, si el componente de la varianza de la formulación de la aleación es grande, entonces esto implica que la variabilidad global de la dureza podría reducirse utilizando únicamente una de las formulaciones de la aleación.

El cálculo de las sumas de cuadrados y el análisis de varianza del diseño anidado de m etapas son similares al análisis presentado en la sección 13-1. Por ejemplo, el análisis de varianza del diseño anidado de tres etapas se resume en la tabla 13-7. En esta tabla se muestran también las definiciones de las sumas de cuadrados. Observe que son una extensión simple de las fórmulas para el diseño anidado de dos etapas. Muchos paquetes de software de estadística realizarán los cálculos.

Para determinar los estadísticos de prueba apropiados deben encontrarse los cuadrados medios esperados empleando los métodos del capítulo 12. Por ejemplo, si los factores A y B son fijos y el factor C es aleatorio, entonces los cuadrados medios esperados pueden derivarse como se indica en la tabla 13-8. En esta tabla se indican los estadísticos de prueba apropiados para esta situación.

Tabla 13-8 Derivación de los cuadrados medios esperados para un diseño anidado de tres etapas con A y B fijos y C aleatorio

Factor	F a i	F b j	R c k	R n l	Cuadrado medio esperado
τ_i	0	b	c	n	$\sigma^2 + n\sigma_\gamma^2 + \frac{bcn\sum \tau_i^2}{a-1}$
$\beta_{j(i)}$	1	0	c	n	$\sigma^2 + n\sigma_\gamma^2 + \frac{cn\sum\sum \beta_{j(i)}^2}{a(b-1)}$
$\gamma_{k(ij)}$	1	1	1	n	$\sigma^2 + n\sigma_\gamma^2$
$\varepsilon_{l(ijk)}$	1	1	1	1	σ^2

13-3 DISEÑOS CON FACTORES ANIDADOS Y FACTORIALES

En experimentos con factores múltiples, algunos factores pueden estar incluidos en un arreglo factorial y otros estar anidados. En ocasiones a estos diseños se les llama **diseños factoriales-anidados.** El análisis estadístico de un diseño así con tres factores se ilustra en el ejemplo siguiente.

EJEMPLO 13-2

Un ingeniero industrial estudia la inserción manual de componentes electrónicos en tarjetas de circuitos impresos a fin de mejorar la rapidez de la operación de ensamblaje. Ha diseñado tres dispositivos de ensamblaje y dos arreglos del sitio de trabajo que parecen prometedores. Se necesitan operadores para realizar el ensamblaje, y se decide seleccionar aleatoriamente cuatro operadores para cada combinación dispositivo-arreglo del sitio de trabajo. Sin embargo, debido a que los sitios de trabajo se encuentran en diferentes puntos dentro de la planta, es difícil usar los *mismos* cuatro operadores para cada arreglo del sitio de trabajo. Por lo tanto, los cuatro operadores escogidos para el arreglo 1 son diferentes de los cuatro para el arreglo 2. Puesto que sólo hay tres dispositivos y dos arreglos del sitio de trabajo, y los operadores se escogen al azar, se trata de un **modelo mixto.** Las combinaciones de tratamientos de este diseño se corren en orden aleatorio y se obtienen dos réplicas. Los tiempos de ensamblaje se miden en segundos y se muestran en la tabla 13-9.

En este experimento, los operadores están anidados dentro de los niveles de los arreglos del sitio de trabajo, mientras que los dispositivos y los arreglos del sitio de trabajo están incluidos en un factorial. Por lo tanto, este diseño tiene factores anidados y factoriales. El **modelo lineal** para este diseño es

$$y_{ijkl} = \mu + \tau_i + \beta_j + \gamma_{k(j)} + (\tau\beta)_{ij} + (\tau\gamma)_{ik(j)} + \varepsilon_{(ijk)l} \quad \begin{cases} i = 1, 2, 3 \\ j = 1, 2 \\ k = 1, 2, 3, 4 \\ l = 1, 2 \end{cases} \tag{13-14}$$

donde τ_i es el efecto del dispositivo i-ésimo, β_j es el efecto del arreglo del sitio de trabajo j-ésimo, $\gamma_{k(j)}$ es el efecto del operador k-ésimo dentro del nivel j-ésimo del arreglo del sitio de trabajo, $(\tau\beta)_{ij}$ es la interacción dispositivo × arreglo del sitio de trabajo, $(\tau\gamma)_{ik(j)}$ es la interacción dispositivo × operadores dentro del arreglo del sitio de trabajo, y $\varepsilon_{(ijk)l}$ es el término del error usual. Observe que no puede existir ninguna interacción arreglo del sitio de trabajo × operador porque no todos los operadores usan todos los arreglos del sitio de trabajo. Asimismo, tampoco puede haber ninguna interacción dispositivo × arreglo del sitio de trabajo × operador. En la tabla 13-10 se derivan los cuadrados medios esperados utilizando el algoritmo tabular del capítulo 12. Esto produce el análisis de un modelo mixto **restringido.** El estadístico de prueba apropiado para cualquier efecto o interacción puede encontrarse inspeccionando esta tabla.

Tabla 13-9 Datos del tiempo de ensamblaje del ejemplo 13-2

	Arreglo 1				Arreglo 2				
Operador	1	2	3	4	1	2	3	4	$y_{i...}$
Dispositivos 1	22	23	28	25	26	27	28	24	404
	24	24	29	23	28	25	25	23	
Dispositivos 2	30	29	30	27	29	30	24	28	447
	27	28	32	25	28	27	23	30	
Dispositivos 3	25	24	27	26	27	26	24	28	401
	21	22	25	23	25	24	27	27	
Totales de los operadores, $y_{.jk.}$	149	150	171	149	163	159	151	160	
Totales de los arreglos, $y_{.j..}$		619				633			$1252 = y_{....}$

Tabla 13-10 Derivación del cuadrado medio esperado del ejemplo 13-2

Factor	*F* 3 *i*	*F* 2 *j*	*R* 4 *k*	*R* 2 *l*	Cuadrado medio esperado
τ_i	0	2	4	2	$\sigma^2 + 2\sigma_{\tau\gamma}^2 + 8\sum \tau_i^2$
β_j	3	0	4	2	$\sigma^2 + 6\sigma_\gamma^2 + 24\sum \beta_j^2$
$\gamma_{k(j)}$	3	1	1	2	$\sigma^2 + 6\sigma_\gamma^2$
$(\tau\beta)_{ij}$	0	0	4	2	$\sigma^2 + 2\sigma_{\tau\gamma}^2 + 4\sum\sum (\tau\beta)_{ij}^2$
$(\tau\gamma)_{ik(j)}$	0	1	1	2	$\sigma^2 + 2\sigma_{\tau\gamma}^2$
$\varepsilon_{(ijk)l}$	1	1	1	1	σ^2

En la tabla 13-11 se muestra el análisis de varianza completo. Se observa que los dispositivos de ensamblaje son significativos y que los operadores dentro de los arreglos del sitio de trabajo también difieren significativamente. Está presente también una interacción significativa entre los dispositivos y los operadores dentro de los arreglos del sitio de trabajo, indicando que los efectos de los diferentes dispositivos no son los mismos para todos los operadores. Los arreglos del sitio de trabajo parecen tener un efecto reducido sobre el tiempo de ensamblaje. Por lo tanto, para minimizar el tiempo de ensamblaje, la atención debería centrarse en los dispositivos tipo 1 y 3. (Observe que los totales de los dispositivos de la tabla 13-9 son menores para los tipos 1 y 3 que para el tipo 2. Esta diferencia en las medias del tipo de dispositivo podría probarse formalmente utilizando comparaciones múltiples.) Además, la interacción entre los operadores y los dispositivos implica que algunos operadores son más eficientes que otros al utilizar los mismos dispositivos. Quizás estos efectos operador-dispositivo podrían aislarse y los operadores cuyo desempeño es menos eficiente podrían mejorar impartiéndoles capacitación adicional.

Cálculos

Hay varios paquetes de software de estadística que analizan con facilidad diseños factoriales-anidados, incluyendo Minitab y SAS. En la tabla 13-12 se presenta la salida de Minitab (Balanced ANOVA, análisis de varianza balanceado), suponiendo la forma restringida del modelo mixto, para el ejemplo 13-2. Los cuadrados medios esperados de la parte inferior de la tabla 13-12 concuerdan con los que se derivaron con el método tabular de la tabla 13-10. *Q*[1], *Q*[3] y *Q*[4] son los efectos del factor fijo para los arreglos del sitio de trabajo, los dispositivos, y la interacción arreglo del sitio de trabajo × dispositivo, respectivamente. Las estimaciones de los componentes de la varianza son:

Operador (arreglo): $\sigma_\gamma^2 = 1.609$

Dispositivo × operador (arreglo): $\sigma_{\tau\gamma}^2 = 1.576$

Error: $\sigma^2 = 2.333$

Tabla 13-11 Análisis de varianza del ejemplo 13-2

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0	Valor *P*
Dispositivos (*F*)	82.80	2	41.40	7.54	0.01
Arreglos (*L*)	4.08	1	4.09	0.34	0.58
Operadores (dentro de los arreglos), *O* (*L*)	71.91	6	11.99	5.15	<0.01
FL	19.04	2	9.52	1.73	0.22
FO(*L*)	65.84	12	5.49	2.36	0.04
Error	56.00	24	2.33		
Total	299.67	47			

Tabla 13-12 Análisis Balanced ANOVA de Minitab del ejemplo 13-2 utilizando el modelo restringido

Análisis de varianza (diseños balanceados)

```
Factor             Type Levels Values
Layout            fixed      2      1      2
Operator(Layout) random      4      1      2      3      4
Fixture           fixed      3      1      2      3

Analysis of Variance for Time

Source                     DF        SS         MS      F      P
Layout                      1     4.083      4.083   0.34  0.581
Operator(Layout)            6    71.917     11.986   5.14  0.002
Fixture                     2    82.792     41.396   7.55  0.008
Layout*Fixture              2    19.042      9.521   1.74  0.218
Fixture*Operator(Layout)   12    65.833      5.486   2.35  0.036
Error                      24    56.000      2.333
Total                      47   299.667

Source                      Variance Error Expected Mean Square for Each Term
                           component term (using restricted model)
 1 Layout                             2    (6) + 6(2) + 24Q[1]
 2 Operator(Layout)            1.609  6    (6) + 6(2)
 3 Fixture                            5    (6) + 2(5) + 16Q[3]
 4 Layout*Fixture                     5    (6) + 2(5) + 8Q[4]
 5 Fixture*Operator(Layout)    1.576  6    (6) + 2(5)
 6 Error                       2.333       (6)
```

Tabla 13-13 Análisis Balanced ANOVA de Minitab del ejemplo 13-2 utilizando el modelo no restringido

Análisis de varianza (diseños balanceados)

```
Factor                 Type Levels Values
Layout                fixed      2      1     2
Operator(Layout)     random      4      1     2     3     4
Fixture               fixed      3      1     2     3

Analysis of Variance for Time

Source                       DF        SS        MS      F      P
Layout                        1     4.083     4.083   0.34  0.581
Operator(Layout)              6    71.917    11.986   2.18  0.117
Fixture                       2    82.792    41.396   7.55  0.008
Layout*Fixture                2    19.042     9.521   1.74  0.218
Fixture*Operator(Layout)     12    65.833     5.486   2.35  0.036
Error                        24    56.000     2.333
Total                        47   299.667

Source                       Variance Error Expected Mean Square for Each Term
                            component  term (using unrestricted model)
  1 Layout                                2  (6) + 2(5) + 6(2) + Q[1,4]
  2 Operator(Layout)            1.083     5  (6) + 2(5) + 6(2)
  3 Fixture                               5  (6) + 2(5) + Q[3,4]
  4 Layout*Fixture                        5  (6) + 2(5) + Q[4]
  5 Fixture*Operator(Layout)    1.576     6  (6) + 2(5)
  6 Error                       2.333        (6)
```

En la tabla 13-13 se presenta el análisis de Minitab del ejemplo 13-2 utilizando la forma **no restringida** del modelo mixto. Los cuadrados medios esperados de la parte inferior de esta tabla son ligeramente diferentes de los que se reportaron para el modelo restringido y, por lo tanto, la construcción de los estadísticos de prueba será ligeramente diferente para el factor operadores (arreglo). Específicamente, el denominador del cociente F de los operadores (arreglo) es la interacción dispositivos × operadores (arreglo) del modelo restringido (12 grados de libertad para el error), y es la interacción arreglo × dispositivos en el modelo no restringido (2 grados de libertad para el error). Puesto que $MS_{\text{arreglo} \times \text{dispositivos}} > MS_{\text{dispositivos} \times \text{operador (arreglo)}}$, y tiene menos grados de libertad, se encuentra ahora que el operador dentro del efecto del arreglo sólo es significativo en el nivel aproximado de 12% (el valor P fue 0.002 en el análisis del modelo restringido). Además, la estimación del componente de la varianza $\hat{\sigma}_{\gamma}^2 = 1.083$ es menor. Sin embargo, puesto que está presente un efecto grande de los dispositivos y una interacción dispositivos × operador (arreglo) significativa, seguiría sospechándose que existe un efecto del operador y, por lo tanto, las conclusiones prácticas de este experimento no son afectadas mucho por elegir la forma restringida o la no restringida del modelo mixto. Las cantidades $Q[1, 4]$ y $Q[3, 4]$ son términos cuadráticos del tipo fijo que contienen el efecto de interacción arreglos × dispositivos.

Si no se cuenta con un paquete de software especializado como SAS o Minitab, entonces puede usarse un programa para analizar experimentos factoriales con factores anidados y factoriales. Así, el experimento del ejemplo 13-2 podría considerarse como un factorial de tres factores, con los dispositivos (F), los operadores (O) y los arreglos (L) como los factores. Entonces se agruparían ciertas sumas de cuadrados y ciertos grados de libertad para formar las cantidades apropiadas requeridas para el diseño con los factores anidados y factoriales de la siguiente manera:

Análisis factorial		Análisis factorial-anidado	
Suma de cuadrados	Grados de libertad	Suma de cuadrados	Grados de libertad
SS_F	2	SS_F	2
SS_L	1	SS_L	1
SS_{FL}	2	SS_{FL}	2
SS_O	3	$SS_{O(L)} = SS_O + SS_{LO}$	6
SS_{LO}	3		
SS_{FO}	6	$SS_{FO(L)} = SS_{FO} + SS_{FOL}$	12
SS_{FOL}	6		
SS_E	24	SS_E	24
SS_T	47	SS_T	47

13-4 DISEÑO DE PARCELAS SUBDIVIDIDAS

En algunos experimentos factoriales con factores múltiples quizá no sea posible la aleatorización completa del orden de las corridas. Esto suele resultar en una generalización del diseño factorial llamada **diseño de parcelas subdivididas.**

Como un ejemplo, considere un fabricante de papel que está interesado en tres métodos diferentes para preparar la pulpa y cuatro temperaturas de cocción diferentes de la pulpa y que desea estudiar el efecto de estos dos factores sobre la resistencia a la tensión del papel. Cada réplica de un experimento factorial requiere 12 observaciones, y el experimentador ha decidido correr tres réplicas. Sin embargo, la ca-

pacidad de la planta piloto sólo permite realizar 12 corridas por día, por lo que el experimentador decide correr una réplica en cada uno de tres días y considerar los días o las réplicas como bloques. En un día lleva a cabo el experimento de la siguiente manera. Se produce un lote de pulpa con uno de los tres métodos bajo estudio. Después este lote se divide en cuatro muestras, y la cocción de cada muestra se hace con una de las cuatro temperaturas. Entonces se produce un segundo lote de pulpa utilizando otro de los tres métodos. Este segundo lote también se divide en cuatro muestras que se prueban con las cuatro temperaturas. Después se repite el proceso, utilizando un lote de pulpa producido por el tercer método. Los datos se muestran en la tabla 13-14.

Inicialmente, esto podría considerarse un experimento factorial con tres niveles del método de preparación (factor A) y cuatro niveles de la temperatura (factor B) en un bloque aleatorizado. Si éste fuera el caso, entonces el orden de experimentación dentro de cada réplica o bloque debería ser completamente aleatorizado. Es decir, dentro de un bloque debería seleccionarse aleatoriamente una combinación de tratamientos (un método de preparación y una temperatura) y obtener una observación, después debería seleccionarse aleatoriamente otra combinación de tratamientos y obtener una segunda observación, y así sucesivamente hasta que se hayan tomado las 12 observaciones en el bloque. Sin embargo, el experimentador no recabó los datos de esta manera. Él hizo un lote de pulpa y obtuvo observaciones para las cuatro temperaturas de ese lote. Debido a la economía para preparar los lotes y al tamaño de los lotes, ésta es la única manera factible de correr este experimento. Un experimento factorial completamente aleatorizado requeriría 36 lotes de pulpa, lo cual es totalmente irrealista. El diseño de parcelas subdivididas requiere sólo tres lotes de pulpa por bloque (réplica), en este caso 9 lotes en total. Evidentemente, el diseño de parcelas subdivididas ha dado como resultado una eficiencia experimental considerable.

El diseño utilizado en el ejemplo de la pulpa es de parcelas subdivididas. Cada réplica o bloque del diseño de parcelas subdivididas se divide en tres partes llamadas **parcelas completas,** y a los métodos de preparación se les llama **tratamientos principales** o **de parcelas completas.** Cada parcela completa se divide en cuatro partes llamadas **subparcelas** (o **parcelas subdivididas**), y se asigna una temperatura a cada una de ellas. A la temperatura se le llama el **tratamiento de la subparcela.** Observe que si están presentes otros factores no controlados o fuera del diseño, y si estos factores no controlados varían cuando los métodos para preparar la pulpa se modifican, entonces cualquier efecto de los factores fuera del diseño sobre la respuesta estará completamente confundido (o mezclado) con el efecto de los métodos para preparar la pulpa. Puesto que los tratamientos de las parcelas completas de un diseño de parcelas subdivididas están confundidos con las parcelas completas y los tratamientos de las subparcelas no están confundidos, es mejor asignar el factor en el que haya mayor interés a las subparcelas, de ser posible.

Este ejemplo es bastante típico de la forma en que se usan los diseños de parcelas subdivididas en un ambiente industrial. Observe que, en esencia, los dos factores "se aplicaron" en tiempos diferentes. Por consiguiente, un diseño de parcelas subdivididas puede considerarse como dos experimentos "combina-

Tabla 13-14 El experimento de la resistencia a la tensión del papel

Método de preparación de la pulpa	Réplica (o bloque) 1			Réplica (o bloque) 2			Réplica (o bloque) 3		
	1	2	3	1	2	3	1	2	3
Temperatura (°F)									
200	30	34	29	28	31	31	31	35	32
225	35	41	26	32	36	30	37	40	34
250	37	38	33	40	42	32	41	39	39
275	36	42	36	41	40	40	40	44	45

dos" o superpuestos entre sí. Un "experimento" tiene el factor parcela completa aplicado a las unidades experimentales grandes (o es un factor cuyos niveles son difíciles de cambiar) y el otro "experimento" tiene el factor subparcela aplicado a las unidades experimentales más pequeñas (o es un factor cuyos niveles son fáciles de cambiar).

El **modelo lineal** para el diseño de parcelas subdivididas es

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \gamma_k + (\tau\gamma)_{ik} + (\beta\gamma)_{jk} + (\tau\beta\gamma)_{ijk} + \varepsilon_{ijk} \qquad \begin{cases} i = 1, 2, \ldots, r \\ j = 1, 2, \ldots, a \\ k = 1, 2, \ldots, b \end{cases} \tag{13-15}$$

donde τ_i, β_j y $(\tau\beta)_{ij}$ representan la parcela completa y corresponden respectivamente a los bloques (o réplicas), a los tratamientos principales (factor A) y al **error de la parcela completa** [réplicas (o bloques) × A); y γ_k, $(\tau\gamma)_{ik}$, $(\beta\gamma)_{jk}$ y $(\tau\beta\gamma)_{ijk}$ representan la subparcela y corresponden respectivamente al tratamiento de la subparcela (factor B), las réplicas (o bloques) × B y las interacciones AB, y al **error de la subparcela** (bloques × AB). Observe que el error de la parcela completa es la interacción réplicas (o bloques) × A y que el error de la subparcela es la interacción de tres factores bloques × AB. Las sumas de cuadrados para estos factores se calculan como en el análisis de varianza de tres factores sin réplicas.

Los cuadrados medios esperados del diseño de parcelas subdivididas, con las réplicas o bloques aleatorios y los tratamientos principales y los tratamientos de subparcelas fijas, se derivan en la tabla 13-15. Observe que el factor principal (A) de la parcela completa se prueba contra el error de la parcela completa, mientras que el subtratamiento (B) se prueba contra la interacción réplicas (o bloques) × subtratamientos. La interacción AB se prueba contra el error de la subparcela. Observe que no hay pruebas para el efecto de la réplica (o bloque) (A) o la interacción réplica (o bloque) × subtratamiento (AC).

El análisis de varianza de los datos de la resistencia a la tensión de la tabla 13-14 se resume en la tabla 13-16. Puesto que tanto los métodos de preparación como las temperaturas son fijos y las réplicas son aleatorias, son aplicables los cuadrados medios esperados de la tabla 13-15. El cuadrado medio de los métodos de preparación se compara con el cuadrado medio del error de la parcela completa, y el cuadra-

Tabla 13-15 Derivación del cuadrado medio esperado del diseño de parcelas subdivididas

	Factor	r R i	a F j	b F k	1 R h	Cuadrado medio esperado
Parcela completa	τ_i	1	a	b	1	$\sigma^2 + ab\sigma_\tau^2$
	β_j	r	0	b	1	$\sigma^2 + b\sigma_{\tau\beta}^2 + \frac{rb\sum \beta_j^2}{a-1}$
	$(\tau\beta)_{ij}$	1	0	b	1	$\sigma^2 + b\sigma_{\tau\beta}^2$
Subparcela	γ_k	r	a	0	1	$\sigma^2 + a\sigma_{\tau\gamma}^2 + \frac{ra\sum \gamma_k^2}{(b-1)}$
	$(\tau\gamma)_{ik}$	1	a	0	1	$\sigma^2 + a\sigma_{\tau\gamma}^2$
	$(\beta\gamma)_{jk}$	r	0	0	1	$\sigma^2 + \sigma_{\tau\beta\gamma}^2 + \frac{r\sum\sum (\beta\gamma)_{jk}^2}{(a-1)(b-1)}$
	$(\tau\beta\gamma)_{ijk}$	1	0	0	1	$\sigma^2 + \sigma_{\tau\beta\gamma}^2$
	$\varepsilon_{(ijk)h}$	1	1	1	1	σ^2 (no estimable)

subtratamiento. Esta situación también es similar a un **submuestreo,** como lo describe Ostle [92]. Suponiendo que A y B son fijos, los cuadrados medios esperados en este caso son

$$
\begin{aligned}
E(MS_A) &= \sigma_\theta^2 + n\sigma_\phi^2 + \frac{bn\sum \tau_i^2}{a-1} \\
E(MS_B) &= \sigma_\theta^2 + n\sigma_\phi^2 + \frac{an\sum \beta_j^2}{b-1} \\
E(MS_{AB}) &= \sigma_\theta^2 + n\sigma_\phi^2 + \frac{n\sum\sum (\tau\beta)_{ij}^2}{(a-1)(b-1)} \\
E(MS_E) &= \sigma_\theta^2
\end{aligned}
\tag{13-18}
$$

Por lo tanto, no hay pruebas para los efectos principales a menos que la interacción sea insignificante. La situación es exactamente la de un análisis de varianza de dos factores con una observación por celda. Si los dos factores son aleatorios, entonces los efectos principales pueden probarse contra la interacción AB. Si sólo uno de los factores es aleatorio, entonces el factor fijo puede probarse contra la interacción AB.

En general, si se analiza un diseño factorial y todos los efectos principales y las interacciones son significativos, entonces deberá examinarse con atención **cómo se realizó realmente el experimento.** Puede haber restricciones sobre la aleatorización en el modelo que no se tomaron en cuenta en el análisis y, por consiguiente, los datos no deberán analizarse como un factorial.

13-5 OTRAS VARIANTES DEL DISEÑO DE PARCELAS SUBDIVIDIDAS

13-5.1 Diseño de parcelas subdivididas con más de dos factores

En ocasiones se encuentra que la parcela completa o la subparcela contendrán dos o más factores, dispuestos en una estructura factorial. Como un ejemplo, considere un experimento conducido en un horno para hacer crecer un óxido en una oblea de silicio. Las variables de respuesta de interés son el espesor de la capa de óxido y la uniformidad de la capa. Hay cuatro factores del diseño: la temperatura (A), el flujo de gas (B), el tiempo (C) y la posición de la oblea en el horno (D). El experimentador planea correr un diseño factorial 2^4 con dos réplicas (32 ensayos). Ahora bien, los factores A y B (la temperatura y el flujo de gas) son difíciles de cambiar, mientras que C y D (el tiempo y la posición de la oblea) son fáciles de modificar. Esto lleva al diseño de parcelas subdivididas que se muestra en la figura 13-7. Observe que las dos réplicas del experimento están subdivididas en cuatro parcelas completas, cada una de las cuales contiene una combinación de los ajustes de la temperatura y el flujo de gas. Una vez que se eligen estos niveles, cada parcela completa se subdivide en cuatro subparcelas, y se realiza un diseño factorial 2^2 en los factores tiempo y posición de la oblea, donde las combinaciones de tratamientos de la subparcela se prueban en orden aleatorio. Únicamente se hacen cuatro cambios de la temperatura y del flujo de gas en cada réplica, mientras que los niveles del tiempo y la posición de la oblea están completamente aleatorizados.

Un modelo para este experimento, consistente con la ecuación 13-16, es

$$
\begin{aligned}
y_{ijklm} = {} & \mu + \tau_i + \beta_j + \gamma_k + (\beta\gamma)_{jk} + \theta_{ijk} + \delta_l + \lambda_m + (\delta\lambda)_{lm} \\
& + (\beta\delta)_{jl} + (\beta\lambda)_{jm} + (\gamma\delta)_{kl} + (\delta\lambda)_{lm} + (\beta\gamma\delta)_{jkl} + (\beta\gamma\lambda)_{jkm} \\
& + (\beta\delta\lambda)_{jlm} + (\gamma\delta\lambda)_{klm} + (\beta\gamma\delta\lambda)_{jklm} + \varepsilon_{ijklm} \qquad \begin{cases} i = 1, 2 \\ j = 1, 2 \\ k = 1, 2 \\ l = 1, 2 \\ m = 1, 2 \end{cases}
\end{aligned}
\tag{13-19}
$$

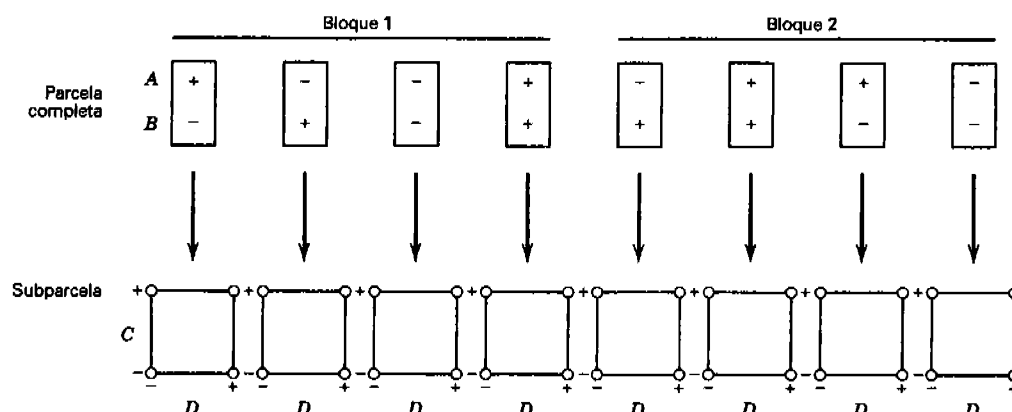


Figura 13-7 Diseño de parcelas subdivididas con cuatro factores del diseño, dos en la parcela completa y dos en la subparcela.

donde τ_i representa el efecto de la réplica, β_j y γ_k los efectos principales de la parcela completa, θ_{ijk} es el error de la parcela completa, δ_l y λ_m representan los efectos principales de la subparcela y ε_{ijklm} es el error de la subparcela. Se han incluido todas las interacciones entre los cuatro factores del diseño. En la tabla 13-17 se presenta el análisis de varianza de este diseño, suponiendo que las réplicas son aleatorias y que todos los factores del diseño son efectos fijos. En esta tabla, σ_θ^2 y σ_ε^2 representan las varianzas de los errores de la parcela completa y de la subparcela, respectivamente, σ_τ^2 es la varianza de los efectos de los blo-

Tabla 13-17 Análisis abreviado de un diseño de parcelas subdivididas con los factores A y B en las parcelas completas y los factores C y D en las subparcelas (referirse a la figura 13-7)

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio esperado
Réplicas (τ_i)	$SS_{\text{Réplicas}}$	1	$\sigma_\varepsilon^2 + 16\sigma_\tau^2$
A (β_j)	SS_A	1	$\sigma_\varepsilon^2 + 8\sigma_\theta^2 + A$
B (γ_k)	SS_B	1	$\sigma_\varepsilon^2 + 8\sigma_\theta^2 + B$
AB	SS_{AB}	1	$\sigma_\varepsilon^2 + 8\sigma_\theta^2 + AB$
Error de la parcela completa (θ_{ijk})	SS_{WP}	3	$\sigma_\varepsilon^2 + 8\sigma_\theta^2$
C (δ_l)	SS_C	1	$\sigma_\varepsilon^2 + C$
D (λ_m)	SS_D	1	$\sigma_\varepsilon^2 + D$
CD	SS_{CD}	1	$\sigma_\varepsilon^2 + CD$
AC	SS_{AC}	1	$\sigma_\varepsilon^2 + AC$
BC	SS_{BC}	1	$\sigma_\varepsilon^2 + BC$
AD	SS_{AD}	1	$\sigma_\varepsilon^2 + AD$
BD	SS_{BD}	1	$\sigma_\varepsilon^2 + BD$
ABC	SS_{ABC}	1	$\sigma_\varepsilon^2 + ABC$
ABD	SS_{ABD}	1	$\sigma_\varepsilon^2 + ABD$
ACD	SS_{ACD}	1	$\sigma_\varepsilon^2 + ACD$
BCD	SS_{BCD}	1	$\sigma_\varepsilon^2 + BCD$
$ABCD$	SS_{ABCD}	1	$\sigma_\varepsilon^2 + ABCD$
Error de la subparcela (ε_{ijklm})	SS_{SP}	12	σ_ε^2
Total	SS_T	31	

ques, y (para simplificar) se han usado letras mayúsculas latinas para denotar los efectos de tipo fijo. Los efectos principales y la interacción de la parcela completa se prueban contra el error de la parcela completa, mientras que los factores de la subparcela y todas las demás interacciones se prueban contra el error de la subparcela. Si algunos de los factores del diseño son aleatorios, los estadísticos de prueba serán diferentes. En algunos casos no habrá ninguna prueba F exacta y deberá usarse el procedimiento de Satterthwaite (descrito en el capítulo 12).

Los experimentos factoriales con tres o más factores en una estructura de parcelas subdivididas tienden a ser experimentos bastante grandes. Por otra parte, la estructura de parcelas subdivididas con frecuencia facilita la realización de un experimento grande. Por ejemplo, en el caso del horno de oxidación, los experimentadores sólo tienen que cambiar ocho veces los factores que son difíciles de modificar (A y B), por lo que quizá un experimento de 32 corridas no sea demasiado ilógico. Es posible reducir el número de corridas utilizando un factorial fraccionado para los factores del diseño de interés.

13-5.2 Diseño de parcelas con doble subdivisión

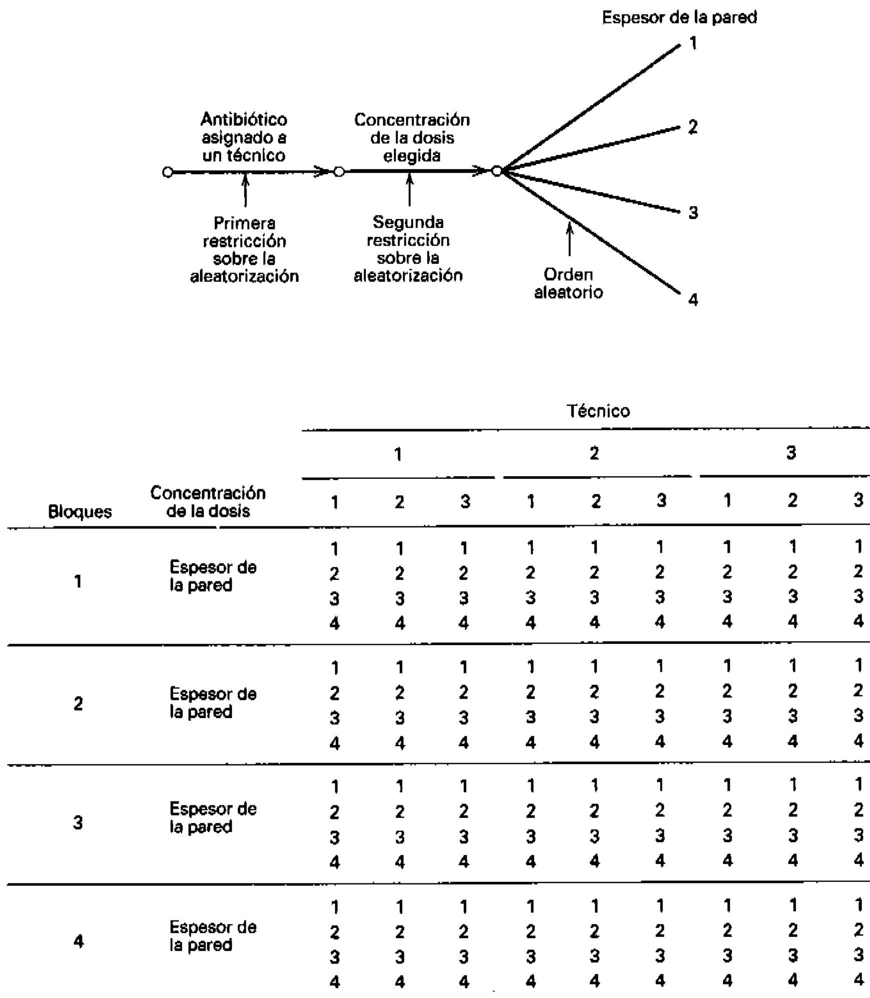
El concepto de diseños de parcelas subdivididas puede extenderse a situaciones en las que pueden ocurrir restricciones sobre la aleatorización en cualquier número de niveles dentro del experimento. Si hay dos niveles de restricciones sobre la aleatorización, al arreglo se le llama **diseño de parcelas con doble subdivisión.** En el ejemplo siguiente se ilustra este diseño.

EJEMPLO 13-3 ··

Un investigador estudia los tiempos de absorción de un tipo particular de cápsula de antibiótico. Hay tres técnicos, tres concentraciones de la dosis y cuatro espesores de la pared de la cápsula. Cada réplica de un experimento factorial requeriría 36 observaciones. El experimentador se ha decidido por cuatro réplicas, y es necesario correr cada réplica en un día diferente. Observe que los días pueden considerarse como bloques. Dentro de una réplica (o un bloque) (día), el experimento se realiza asignando una unidad de antibiótico a un técnico que lleva a cabo el experimento con las tres concentraciones de la dosis y los cuatro espesores de la pared. Una vez que se ha formulado una concentración particular de la dosis, se prueban los cuatro espesores de la pared con esa concentración. Después se selecciona otra concentración de la dosis y se prueban los cuatro espesores de la pared. Por último se prueba la tercera concentración de la dosis y los cuatro espesores de la pared. Mientras tanto, otros dos técnicos del laboratorio también siguen el mismo plan, empezando cada uno con una unidad de antibiótico.

Observe que hay *dos* restricciones sobre la aleatorización dentro de cada réplica (o bloque): el técnico y la concentración de la dosis. Las parcelas completas corresponden al técnico. El orden en que se asignan los técnicos a las unidades de antibiótico se determina aleatoriamente. Las concentraciones de la dosis forman tres subparcelas. La concentración de la dosis puede asignarse aleatoriamente a una subparcela. Por último, dentro de una concentración particular de la dosis se prueban los cuatro espesores de la pared de la cápsula de manera aleatoria, formando cuatro sub-subparcelas. A los espesores de la pared suele llamárseles sub-subtratamientos. Puesto que hay *dos* restricciones sobre la aleatorización en el experimento (algunos autores dicen que hay dos "divisiones" en el diseño), al diseño se le llama diseño de parcelas con doble subdivisión. En la figura 13-8 se ilustran las restricciones sobre la aleatorización y el arreglo experimental de este diseño.

··



		Técnico								
		1			2			3		
Bloques	Concentración de la dosis	1	2	3	1	2	3	1	2	3
1	Espesor de la pared	1	1	1	1	1	1	1	1	1
		2	2	2	2	2	2	2	2	2
		3	3	3	3	3	3	3	3	3
		4	4	4	4	4	4	4	4	4
2	Espesor de la pared	1	1	1	1	1	1	1	1	1
		2	2	2	2	2	2	2	2	2
		3	3	3	3	3	3	3	3	3
		4	4	4	4	4	4	4	4	4
3	Espesor de la pared	1	1	1	1	1	1	1	1	1
		2	2	2	2	2	2	2	2	2
		3	3	3	3	3	3	3	3	3
		4	4	4	4	4	4	4	4	4
4	Espesor de la pared	1	1	1	1	1	1	1	1	1
		2	2	2	2	2	2	2	2	2
		3	3	3	3	3	3	3	3	3
		4	4	4	4	4	4	4	4	4

Figura 13-8 Diseño de parcelas con doble subdivisión.

Un modelo estadístico lineal para el diseño de parcelas con doble subdivisión es

$$\begin{aligned} y_{ijkh} = {} & \mu+\tau_i+\beta_j+(\tau\beta)_{ij}+\gamma_k+(\tau\gamma)_{ik}+(\beta\gamma)_{jk}+(\tau\beta\gamma)_{ijk} \\ & +\delta_h+(\tau\delta)_{ih}+(\beta\delta)_{jh} \\ & +(\tau\beta\delta)_{ijh}+(\gamma\delta)_{kh}+(\tau\gamma\delta)_{ikh}+(\beta\gamma\delta)_{jkh} \\ & +(\tau\beta\gamma\delta)_{ijkh}+\varepsilon_{ijkh} \qquad \begin{cases} i=1,\ 2,\ \ldots,\ r \\ j=1,\ 2,\ \ldots,\ a \\ k=1,\ 2,\ \ldots,\ b \\ h=1,\ 2,\ \ldots,\ c \end{cases} \end{aligned} \tag{13-20}$$

donde τ_i, β_j y $(\tau\beta)_{ij}$ representan la parcela completa y corresponden a las réplicas o bloques, a los tratamientos principales (factor A) y al error de la parcela completa [réplicas (o bloques) × A)], respectiva-

mente; y γ_k, $(\tau\gamma)_{ik}$, $(\beta\gamma)_{jk}$ y $(\tau\beta\gamma)_{ijk}$ representan la subparcela y corresponden al tratamiento de la subparcela (factor B), las interacciones de réplicas (o bloques) $\times$ B y AB, y al error de la subparcela, respectivamente; y δ_h y los parámetros restantes corresponden a la sub-subparcela y representan, respectivamente, al tratamiento de la sub-subparcela (factor C) y a las interacciones restantes. A la interacción de cuatro factores $(\tau\beta\gamma\delta)_{ijkh}$ se le llama el error de la sub-subparcela.

Suponiendo que las réplicas (bloques) son aleatorias y que los demás factores del diseño son fijos, pueden derivarse los cuadrados medios esperados como se muestra en la tabla 13-18. Las pruebas de los tratamientos principales, los subtratamientos, los sub-subtratamientos y sus interacciones son obvias al inspeccionar esta tabla. Observe que no existen pruebas para las réplicas o bloques ni para las interacciones en las que intervienen réplicas o bloques.

El análisis estadístico de un diseño de parcelas con doble subdivisión es como el de una sola réplica de un factorial de cuatro factores. El número de grados de libertad de cada prueba se determina de la manera usual. Para ilustrar, en el ejemplo 13-3, donde se manejaron cuatro réplicas, tres técnicos, tres concentraciones de la dosis y cuatro espesores de la pared, se tendrían sólo $(r-1)(a-1) = (4-1)(3-1) = 6$ grados de libertad del error de la parcela completa para probar a los técnicos. Se trata de un número relativamente pequeño de grados de libertad, y el experimentador podría considerar el uso de réplicas adicio-

Tabla 13-18 Derivación del cuadrado medio esperado para el diseño de parcelas con doble subdivisión

		r	a	b	c	1	
		R	F	F	F	R	
	Factor	i	j	k	h	l	Cuadrado medio esperado
	τ_i	1	a	b	c	1	$\sigma^2 + abcd\sigma_\tau^2$
Parcela completa	β_j	r	0	b	c	1	$\sigma^2 + \sigma_{\tau\beta}^2 + \frac{rbc\sum \beta_j^2}{(a-1)}$
	$(\tau\beta)_{ij}$	1	0	b	c	1	$\sigma^2 + bc\sigma_{\tau\beta}^2$
	γ_k	r	a	0	c	1	$\sigma^2 + ac\sigma_{\tau\gamma}^2 + \frac{rac\sum \gamma_k^2}{(b-1)}$
Subparcela	$(\tau\gamma)_{ik}$	1	a	0	c	1	$\sigma^2 + ac\sigma_{\tau\gamma}^2$
	$(\beta\gamma)_{jk}$	r	0	0	c	1	$\sigma^2 + c\sigma_{\tau\beta\gamma}^2 + \frac{rc\sum\sum (\beta\gamma)_{jh}^2}{(a-1)(b-1)}$
	$(\tau\beta\gamma)_{ijk}$	1	0	0	c	1	$\sigma^2 + c\sigma_{\tau\beta\gamma}^2$
	δ_h	r	a	b	0	1	$\sigma^2 + ab\sigma_{\tau\delta}^2 + \frac{rab\sum \delta_k^2}{(c-1)}$
	$(\tau\delta)_{ih}$	1	a	b	0	1	$\sigma^2 + ab\sigma_{\tau\delta}^2$
	$(\beta\delta)_{jh}$	r	0	b	0	1	$\sigma^2 + b\sigma_{\tau\beta\delta}^2 + \frac{rb\sum\sum (\beta\delta)_{jh}^2}{(a-1)(c-1)}$
	$(\tau\beta\delta)_{ijh}$	1	0	b	0	1	$\sigma^2 + b\sigma_{\tau\beta\delta}^2$
Sub-subparcela	$(\gamma\delta)_{kh}$	r	a	0	0	1	$\sigma^2 + a\sigma_{\tau\gamma\delta}^2 + \frac{ra\sum\sum (\gamma\delta)_{kh}^2}{(b-1)(c-1)}$
	$(\tau\gamma\delta)_{ikh}$	1	a	0	0	1	$\sigma^2 + a\sigma_{\tau\gamma\delta}^2$
	$(\beta\gamma\delta)_{jkh}$	r	0	0	0	1	$\sigma^2 + \sigma_{\tau\beta\gamma\delta}^2 + \frac{r\sum\sum\sum (\beta\gamma\delta)_{ijk}^2}{(b-1)(c-1)}$
	$(\tau\beta\gamma\delta)_{ijkh}$	1	0	0	0	1	$\sigma^2 + \sigma_{\tau\beta\gamma\delta}^2$
	$\varepsilon_{l(ijkh)}$	1	1	1	1	1	σ^2 (no estimable)

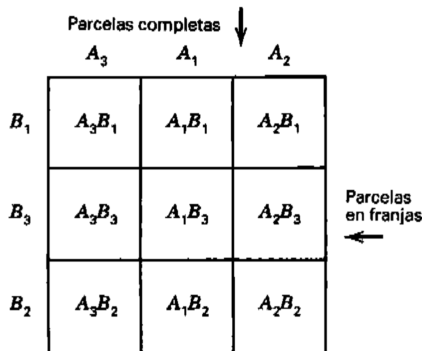


Figura 13-9 Una réplica (bloque) de un diseño de parcelas subdivididas en franjas.

nales para incrementar la precisión de la prueba. Si hay a réplicas, se tendrán $2(r-1)$ grados de libertad para el error de la parcela completa. Por lo tanto, cinco réplicas producirán $2(5-1) = 8$ grados de libertad, seis réplicas producirán $2(6-1) = 10$ grados de libertad, siete réplicas producirán $2(7-1) = 12$ grados de libertad, etcétera. Por consiguiente, es probable que el experimentador no quiera correr menos de cuatro réplicas, ya que se producirían así sólo cuatro grados de libertad. Cada réplica adicional permite ganar dos grados de libertad para el error. Si se cuenta con recursos para correr cinco réplicas, la precisión de la prueba podría incrementarse en un tercio (de seis a ocho grados de libertad). Además, al pasar de cinco a seis réplicas, hay 25% de ganancia adicional en la precisión. Si los recursos lo permiten, el experimentador deberá correr cinco o seis réplicas.

13-5.3 Diseño de parcelas subdivididas en franjas

El diseño de parcelas subdivididas en franjas ha tenido una amplia aplicación en las ciencias agrícolas, pero sólo ocasionalmente encuentra un uso en la experimentación industrial. En el caso más simple, se tienen dos factores A y B. El factor A se aplica a las parcelas completas como en el diseño de parcelas subdivididas estándar. Después el factor B se aplica a **franjas** (que son en realidad sólo otro conjunto de par-

Tabla 13-19 Análisis de varianza abreviado de un diseño de parcelas subdivididas en franjas

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio esperado
Réplicas (o bloques)	$SS_{\text{Réplicas}}$	$r-1$	$\sigma_\varepsilon^2 + ab\sigma_\tau^2$
A	SS_A	$a-1$	$\sigma_\varepsilon^2 + b\sigma_{\tau\beta}^2 + \frac{rb\sum \beta_j^2}{a-1}$
Error_A de la parcela completa	SS_{WP_A}	$(r-1)(a-1)$	$\sigma_\varepsilon^2 + b\sigma_{\tau\beta}^2$
B	SS_B	$b-1$	$\sigma_\varepsilon^2 + a\sigma_{\tau\gamma}^2 + \frac{ra\sum \gamma_k^2}{b-1}$
Error_B de la parcela completa	SS_{WP_B}	$(r-1)(b-1)$	$\sigma_\varepsilon^2 + a\sigma_{\tau\gamma}^2$
AB	SS_{AB}	$(a-1)(b-1)$	$\sigma_\varepsilon^2 + \frac{r\sum\sum (\tau\beta)_{jk}^2}{(a-1)(b-1)}$
Error de la subparcela	SS_{SP}	$(r-1)(a-1)(b-1)$	σ_ε^2
Total	SS_T	$rab-1$	

celas completas) que son ortogonales a las parcelas completas originales utilizadas para el factor A. En la figura 13-9 se ilustra una situación en la que los dos factores A y B tienen tres niveles. Observe que los niveles del factor A están confundidos (o mezclados) con las parcelas completas, y que los niveles del factor B están confundidos con las franjas (las cuales pueden considerarse como un **segundo** conjunto de parcelas completas).

Un modelo para el diseño de parcelas subdivididas en franjas de la figura 13-9, suponiendo γ réplicas, a niveles del factor A y b niveles del factor B, es

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \gamma_k + (\tau\gamma)_{ik} + (\beta\gamma)_{jk} + \varepsilon_{ijk} \qquad \begin{cases} i = 1, 2, \ldots, r \\ j = 1, 2, \ldots, a \\ k = 1, 2, \ldots, b \end{cases}$$

donde $(\tau\beta)_{ij}$ y $(\tau\gamma)_{ik}$ son los errores de la parcela completa de los factores A y B, respectivamente, y ε_{ijk} es el error "de la subparcela" usado para probar la interacción AB. En la tabla 13-19 se muestra un análisis de varianza abreviado, suponiendo que A y B son factores fijos y que las réplicas son aleatorias. En ocasiones las réplicas se consideran como bloques.

13-6 PROBLEMAS

13-1. El fabricante de la carga propulsora de una turbina está estudiando la rapidez de combustión del propulsor obtenido de tres procesos de producción. Se seleccionan al azar cuatro lotes del propulsor de la salida de cada proceso y se hacen tres determinaciones de la rapidez de combustión de cada lote. Los resultados se presentan a continuación. Analizar los datos y sacar conclusiones.

	Proceso 1				Proceso 2				Proceso 3			
Lote	1	2	3	4	1	2	3	4	1	2	3	4
	25	19	15	15	19	23	18	35	14	35	38	25
	30	28	17	16	17	24	21	27	15	21	54	29
	26	20	14	13	14	21	17	25	20	24	50	33

13-2. Se estudia el acabado superficial de piezas metálicas fabricadas en cuatro máquinas. Se conduce un experimento en el que cada máquina es operada por tres operadores diferentes y se colectan y prueban dos ejemplares de cada operador. Debido a la ubicación de las máquinas, se usan operadores diferentes en cada máquina, y los operadores se eligen al azar. Los datos se muestran en la tabla siguiente. Analizar los datos y sacar conclusiones.

	Máquina 1			Máquina 2			Máquina 3			Máquina 4		
Operador	1	2	3	1	2	3	1	2	3	1	2	3
	79	94	46	92	85	76	88	53	46	36	40	62
	62	74	57	99	79	68	75	56	57	53	56	47

13-3. Un ingeniero de manufactura está estudiando la variabilidad dimensional de un componente particular que se produce en tres máquinas. Cada máquina tiene dos mandriles, y se seleccionan al azar cuatro componentes de cada mandril. Los resultados se presentan a continuación. Analizar los datos, suponiendo que las máquinas y los mandriles son factores fijos.

	Máquina 1		Máquina 2		Máquina 3	
Mandril	1	2	1	2	1	2
	12	8	14	12	14	16
	9	9	15	10	10	15
	11	10	13	11	12	15
	12	8	14	13	11	14

13-4. Para simplificar la programación de la producción, un ingeniero industrial está estudiando la posibilidad de asignar un tiempo estándar a una clase particular de tareas, con la creencia de que las diferencias entre las tareas son insignificantes. Para ver si esta simplificación es posible, se seleccionan seis tareas al azar. Cada tarea se encarga a un grupo diferente de tres operadores. Cada operador completa dos veces la tarea en momentos diferentes durante la semana, y se obtienen los resultados siguientes. ¿Qué conclusiones pueden sacarse acerca del uso del tiempo estándar común para todas las tareas de esta clase? ¿Qué valor se usaría para el estándar?

Tarea	Operador 1		Operador 2		Operador 3	
1	158.3	159.4	159.2	159.6	158.9	157.8
2	154.6	154.9	157.7	156.8	154.8	156.3
3	162.5	162.6	161.0	158.9	160.5	159.5
4	160.0	158.7	157.5	158.9	161.1	158.5
5	156.3	158.1	158.3	156.9	157.7	156.9
6	163.7	161.0	162.3	160.3	162.6	161.8

13-5. Considere el diseño anidado de tres etapas que se muestra en la figura 13-5 para investigar la dureza de una aleación. Utilizando los datos que se presentan a continuación, analizar el diseño, suponiendo que la química de la aleación y las hornadas son factores fijos y que los lingotes son aleatorios. Usar la forma restringida del modelo mixto.

	Química de la aleación											
Hornadas	1						2					
	1		2		3		1		2		3	
Lingotes	1	2	1	2	1	2	1	2	1	2	1	2
	40	27	95	69	65	78	22	23	83	75	61	35
	63	30	67	47	54	45	10	39	62	64	77	42

13-6. Analizar nuevamente el experimento del problema 13-5 utilizando la forma no restringida del modelo mixto. Comentar las diferencias que se observan entre los resultados del modelo restringido y el no restringido. Puede usarse un paquete de software de computadora.

13-7. Derivar los cuadrados medios esperados para el diseño anidado balanceado de tres etapas, suponiendo que A es fijo y que B y C son aleatorios. Obtener las fórmulas para estimar los componentes de la varianza. Suponer la forma restringida del modelo mixto.

13-8. Repetir el problema 13-7 suponiendo la forma no restringida del modelo mixto. Puede usarse un paquete de software de computadora para hacerlo. Comentar las diferencias entre el análisis y las conclusiones del modelo restringido y el no restringido.

13-9. Derivar los cuadrados medios esperados para el diseño anidado balanceado de tres etapas si los tres factores son aleatorios. Obtener las fórmulas para estimar los componentes de la varianza.

13-10. Verificar los cuadrados medios esperados que se dan en la tabla 13-1.

13-11. ***Diseños anidados no balanceados.*** Considere un diseño anidado de dos etapas no balanceado con b_j niveles de B bajo el nivel i-ésimo de A y n_{ij} réplicas en la celda ij-ésima.

a) Anotar las ecuaciones normales de mínimos cuadrados para esta situación. Resolver las ecuaciones normales.

b) Construir la tabla del análisis de varianza para el diseño anidado no balanceado de dos etapas.

c) Analizar los datos siguientes, utilizando los resultados del inciso b.

Factor A	1		2		
Factor B	1	2	1	2	3
	6	−3	5	2	1
	4	1	7	4	0
	8		9	3	−3
			6		

13-12. ***Componentes de la varianza en el diseño anidado no balanceado de dos etapas.*** Considere el modelo

$$y_{ijk} = \mu + \tau_i + \beta_{j(i)} + \varepsilon_{k(ij)} \qquad \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, b \\ k = 1, 2, \ldots, n_{ij} \end{cases}$$

donde A y B son factores aleatorios. Demostrar que

$$E(MS_A) = \sigma^2 + c_1\sigma_\beta^2 + c_2\sigma_\tau^2$$
$$E(MS_{B(A)}) = \sigma^2 + c_0\sigma_\beta^2$$
$$E(MS_E) = \sigma^2$$

donde

$$c_0 = \frac{N - \sum_{i=1}^{a}\left(\sum_{j=1}^{b_i} n_{ij}^2 / n_{i.}\right)}{b-a}$$

$$c_1 = \frac{\sum_{i=1}^{a}\left(\sum_{j=1}^{b_i} n_{ij}^2 / n_{i.}\right) - \sum_{i=1}^{a}\sum_{j=1}^{b_i} n_{ij}^2 / N}{a-1}$$

$$c_2 = \frac{N - \dfrac{\sum_{i=1}^{a} n_{i.}^2}{N}}{a-1}$$

13-13. Un ingeniero de procesos está probando el rendimiento de un producto manufacturado en tres máquinas. Cada máquina puede operarse con dos ajustes de la potencia. Además, una máquina tiene tres estaciones en las que se fabrica el producto. Se conduce un experimento en el que cada máquina se prueba con ambos ajustes de la potencia, y se toman tres observaciones del rendimiento de cada estación. Las corridas se hacen en orden aleatorio, y los resultados se presentan a continuación. Analizar este experimento, suponiendo que los tres factores son fijos.

Estación	Máquina 1			Máquina 2			Máquina 3		
	1	2	3	1	2	3	1	2	3
	34.1	33.7	36.2	31.1	33.1	32.8	32.9	33.8	33.6
Ajuste de la potencia 1	30.3	34.9	36.8	33.5	34.7	35.1	33.0	33.4	32.8
	31.6	35.0	37.1	34.0	33.9	34.3	33.1	32.8	31.7
	24.3	28.1	25.7	24.1	24.1	26.0	24.2	23.2	24.7
Ajuste de la potencia 2	26.3	29.3	26.1	25.0	25.1	27.1	26.1	27.4	22.0
	27.1	28.6	24.9	26.3	27.9	23.9	25.3	28.0	24.8

13-14. Suponga que en el problema 13-13 podrían emplearse un gran número de ajustes de la potencia y que los dos que se seleccionaron para el experimento se escogieron al azar. Obtener los cuadrados medios esperados para esta situación suponiendo la forma restringida del modelo mixto y hacer las modificaciones apropiadas al análisis anterior.

13-15. Analizar nuevamente el experimento del problema 13-14 suponiendo la forma no restringida del modelo mixto. Puede usarse un paquete de software de computadora para hacerlo. Comentar las diferencias entre el análisis y las conclusiones del modelo restringido y el no restringido.

13-16. Un ingeniero de estructuras está estudiando la resistencia de una aleación de aluminio adquirida de tres fabricantes. Cada fabricante entrega la aleación en barras de tamaño estándar de 1.0, 1.5 o 2.0 pulgadas. El procesamiento de los diferentes tamaños de las barras a partir de un lingote común implica técnicas diferentes de forjado, por lo que este factor puede ser importante. Además, las barras se forjan de lingotes fabricados en hornadas diferentes. Cada fabricante entrega dos ejemplares de prueba de cada tamaño de las barras de tres hornadas. Los datos de la resistencia resultantes se presentan a continuación. Analizar los datos, suponiendo que los fabricantes y el tamaño de las barras son fijos y las hornadas son aleatorias. Usar la forma restringida del modelo mixto.

Hornada	Fabricante 1			Fabricante 2			Fabricante 3		
	1	2	3	1	2	3	1	2	3
Tamaño de la barra: 1 pulgada	1.230	1.346	1.235	1.301	1.346	1.315	1.247	1.275	1.324
	1.259	1.400	1.206	1.263	1.392	1.320	1.296	1.268	1.315
$1\frac{1}{2}$ pulgadas	1.316	1.329	1.250	1.274	1.384	1.346	1.273	1.260	1.392
	1.300	1.362	1.239	1.268	1.375	1.357	1.264	1.265	1.364
2 pulgadas	1.287	1.346	1.273	1.247	1.362	1.336	1.301	1.280	1.319
	1.292	1.382	1.215	1.215	1.328	1.342	1.262	1.271	1.323

13-17. Resolver de nuevo el problema 13-16 utilizando la forma no restringida del modelo mixto. Puede usarse un paquete de software de computadora para hacerlo. Comentar cualquier diferencia entre el análisis y las conclusiones del modelo restringido y el no restringido.

13-18. Suponga que en el problema 13-16 las barras pueden adquirirse en muchos tamaños y que los tres tamaños que realmente se utilizaron en el experimento fueron seleccionados al azar. Obtener los cuadrados medios esperados para esta situación y hacer las modificaciones apropiadas al análisis anterior. Usar la forma restringida del modelo mixto.

13-19. La normalización del acero se hace calentándolo arriba de la temperatura crítica, recalentándolo y después enfriándolo con aire. Este proceso incrementa la resistencia del acero, refina el grano y homogeneiza la estructura. Se lleva a cabo un experimento para determinar el efecto de la temperatura y de la duración del tratamiento térmico sobre la resistencia del acero normalizado. Se seleccionan dos temperaturas y tres dura-

ciones. El experimento se realiza calentando el horno a una temperatura seleccionada aleatoriamente e insertando tres ejemplares de prueba. Después de 10 minutos se retira uno de ellos, después de 20 minutos se retira un segundo ejemplar y después de 30 minutos se retira el último. Entonces se corre la temperatura al otro nivel y se repite el proceso. Se requieren cuatro corrimientos para recabar los datos, los cuales se muestran abajo. Analizar los datos y sacar conclusiones, suponiendo que ambos factores son fijos.

		Temperatura, °F	
Corrimiento	Tiempo, minutos	1500	1600
	10	63	89
1	20	54	91
	30	61	62
	10	50	80
2	20	52	72
	30	59	69
	10	48	73
3	20	74	81
	30	71	69
	10	54	88
4	20	48	92
	30	59	64

13-20. Se diseña un experimento para estudiar la dispersión de los pigmentos de una pintura. Se estudian cuatro mezclas diferentes de un pigmento particular. El procedimiento consiste en preparar una mezcla particular y en aplicarla después a un panel utilizando tres métodos (con brocha, por rocío y con rodillo). La respuesta medida es el porcentaje de reflectancia (coeficiente de reflexión) del pigmento. Se necesitan tres días para correr el experimento, y los datos obtenidos se presentan a continuación. Analizar los datos y sacar conclusiones, suponiendo que las mezclas y los métodos de aplicación son fijos.

		Mezcla			
Día	Método de aplicación	1	2	3	4
	1	64.5	66.3	74.1	66.5
1	2	68.3	69.5	73.8	70.0
	3	70.3	73.1	78.0	72.3
	1	65.2	65.0	73.8	64.8
2	2	69.2	70.3	74.5	68.3
	3	71.2	72.8	79.1	71.5
	1	66.2	66.5	72.3	67.7
3	2	69.0	69.0	75.4	68.6
	3	70.8	74.2	80.1	72.4

13-21. Repetir el problema 13-20, suponiendo que las mezclas son aleatorias y que los métodos de aplicación son fijos.

13-22. Considere el diseño de parcelas con doble subdivisión del ejemplo 13-3. Suponga que este experimento se conduce como se describe y que se obtienen los datos que se muestran en la siguiente tabla. Analizar los datos y sacar conclusiones.

Réplicas (o bloques)	Concentración de las dosis	Técnico 1			Técnico 2			Técnico 3		
		1	2	3	1	2	3	1	2	3
	Espesor de la pared									
1	1	95	71	108	96	70	108	95	70	100
	2	104	82	115	99	84	100	102	81	106
	3	101	85	117	95	83	105	105	84	113
	4	108	85	116	97	85	109	107	87	115
2	1	95	78	110	100	72	104	92	69	101
	2	106	84	109	101	79	102	100	76	104
	3	103	86	116	99	80	108	101	80	109
	4	109	84	110	112	86	109	108	86	113
3	1	96	70	107	94	66	100	90	73	98
	2	105	81	106	100	84	101	97	75	100
	3	106	88	112	104	87	109	100	82	104
	4	113	90	117	121	90	117	110	91	112
4	1	90	68	109	98	68	106	98	72	101
	2	100	84	112	102	81	103	102	78	105
	3	102	85	115	100	85	110	105	80	110
	4	114	88	118	118	85	116	110	95	120

13-23. Resolver nuevamente el problema 13-22, suponiendo que los técnicos se eligen al azar. Usar la forma restringida del modelo mixto.

13-24. Suponga que en el problema 13-22 se usaron cuatro técnicos. Suponiendo que todos los factores son fijos, ¿cuántos bloques deberán correrse para obtener un número adecuado de grados de libertad para probar las diferencias entre los técnicos?

13-25. Considere el experimento que se describe en el ejemplo 13-3. Demostrar cómo se determinaría el orden en que se corren la combinaciones de tratamientos si este experimento se realizara como a) una parcela con doble subdivisión, b) una parcela subdividida, c) un diseño factorial en un bloque aleatorizado y d) un diseño factorial completamente aleatorizado.

14 Otros tópicos de diseño y análisis

El tema de los experimentos diseñados estadísticamente es muy amplio. En los capítulos previos se ha ofrecido una presentación introductoria de muchos de los conceptos y métodos básicos, aunque en algunos casos sólo se ha podido presentar un panorama general. Por ejemplo, hay exposiciones que ocupan un libro sobre tópicos, como la metodología de superficies de respuesta, los experimentos con mezclas, la estimación de los componentes de la varianza y los diseños óptimos. En este capítulo se presenta un panorama general de varios tópicos más que el experimentador puede encontrar potencialmente útiles.

14-1 RESPUESTAS Y TRANSFORMACIONES NO NORMALES

14-1.1 Selección de una transformación: el método de Box-Cox

En la sección 3-4.3 se estudió el problema de una varianza no constante de la variable de respuesta y en un experimento diseñado, y se señaló que se trata de una desviación de los supuestos del análisis de varianza estándar. Este problema de la desigualdad de la varianza ocurre con relativa frecuencia en la práctica, muchas veces en conjunción con una variable de respuesta no normal. Algunos ejemplos incluirían el conteo de defectos o partículas, los datos de proporciones, como el rendimiento o la proporción de productos defectuosos, o una variable de respuesta que sigue alguna distribución sesgada (una "cola" de la distribución de la respuesta es más larga que la otra). Se introdujo la **transformación de la variable de respuesta** como un método apropiado para estabilizar la varianza de la respuesta. Se revisaron dos métodos para seleccionar la **forma** de la transformación, una técnica gráfica empírica y un procedimiento esencialmente de ensayo y error en el que el experimentador simplemente intenta una o más transformaciones, y selecciona la que produce la gráfica más agradable o satisfactoria de los residuales contra la respuesta ajustada.

En general, las transformaciones se usan para tres propósitos: estabilizar la varianza de la respuesta, hacer que la distribución de la variable de respuesta esté más cerca de la distribución normal y mejorar el ajuste del modelo a los datos. Este último objetivo podría incluir la simplificación del modelo, por ejem-

plo, eliminando términos de interacción. En ocasiones, una transformación será razonablemente eficaz para conseguir de manera simultánea más de uno de estos objetivos.

Se ha señalado ya que la **familia de potencias** de las transformaciones $y^* = y^\lambda$ es muy útil, donde λ es el parámetro de la transformación que habrá de determinarse (por ejemplo $\lambda = \frac{1}{2}$ significa usar la raíz cuadrada de la respuesta original). Box y Cox [15] han indicado cómo puede estimarse el parámetro de la transformación λ al mismo tiempo que los demás parámetros del modelo (la media global y los efectos de los tratamientos). La teoría fundamental en su procedimiento utiliza el método de máxima verosimilitud. El procedimiento de cálculo real consiste en efectuar, para varios valores de λ, el análisis de varianza estándar de

$$y^{(\lambda)} = \begin{cases} \dfrac{y^\lambda - 1}{\lambda \dot{y}^{\lambda-1}} & \lambda \neq 0 \\ \dot{y} \ln y & \lambda = 0 \end{cases} \tag{14-1}$$

donde $\dot{y} = \ln^{-1}[(1/n) \Sigma \ln y]$ es la media geométrica de las observaciones. La estimación de máxima verosimilitud de λ es el valor para el que la suma de cuadrados del error, por ejemplo $SS_E(\lambda)$, es un mínimo. Este valor de λ se encuentra generalmente construyendo una gráfica de $SS_E(\lambda)$ contra λ y leyendo después en la gráfica el valor de λ que minimiza $SS_E(\lambda)$. En general, son suficientes entre 10 y 20 valores de λ para estimar el valor óptimo. Si se necesita una estimación más precisa de λ, podría realizarse una segunda iteración utilizando un número mayor de valores.

Observe que *no es posible* seleccionar el valor de λ comparando *directamente* las sumas de cuadrados del error obtenidas en los análisis de varianza de y^λ, ya que para cada valor de λ la suma de cuadrados del error se mide en una escala diferente. Además, surge un problema con y cuando $\lambda = 0$; a saber, cuando λ tiende a cero, y^λ tiende a la unidad. Es decir, cuando $\lambda = 0$, todos los valores de la respuesta son una constante. El componente $(y^\lambda - 1)/\lambda$ de la ecuación 14-1 alivia este problema porque cuando λ tiende a cero, $(y^\lambda - 1)/\lambda$ tiende a un límite de $\ln y$. El componente del divisor $\dot{y}^\lambda - 1$ de la ecuación 14-1 reescala las respuestas para que las sumas de cuadrados del error sean comparables directamente.

Al utilizar el método de Box-Cox, se recomienda que el experimentador use elecciones simples de λ, ya que es probable que la diferencia práctica entre $\lambda = 0.5$ y $\lambda = 0.58$ sea pequeña, pero la transformación de la raíz cuadrada ($\lambda = 0.5$) es mucho más fácil de interpretar. Obviamente, los valores de λ próximos a la unidad sugerirían que no es necesaria ninguna transformación.

Una vez que se ha seleccionado un valor de λ por el método de Box-Cox, el experimentador puede analizar los datos utilizando y^λ como la respuesta, a menos desde luego que $\lambda = 0$, en cuyo caso se usa $\ln y$. Es perfectamente aceptable utilizar $y^{(\lambda)}$ como la respuesta real, aun cuando las estimaciones de los parámetros del modelo tendrán una diferencia de escala y un corrimiento del origen en comparación con los resultados obtenidos cuando se usa y^λ (o $\ln y$).

Es posible encontrar un intervalo de confianza aproximado de $100(1 - \alpha)$ por ciento para λ calculando

$$SS^* = SS_E(\lambda)\left(1 + \frac{t^2_{\alpha/2,\nu}}{\nu}\right) \tag{14-2}$$

donde ν es el número de grados de libertad, y graficando una recta paralela al eje λ a la altura SS^* sobre la gráfica de $SS_E(\lambda)$ contra λ. Entonces, al localizar los puntos sobre el eje λ donde SS^* corta la curva $SS_E(\lambda)$, pueden leerse directamente en la gráfica los límites de confianza para λ. Si este intervalo de confianza incluye el valor $\lambda = 1$, esto implica (como se señaló antes) que los datos no soportan la necesidad de una transformación.

EJEMPLO 14-1 ....

El procedimiento de Box-Cox se ilustrará utilizando los datos de la descarga pico presentados originalmente en el ejemplo 3-5. Recuerde que se trata de un experimento con un solo factor (ver la tabla 3-7 para los datos originales). Utilizando la ecuación 14-1 se calcularon los valores de $SS_E(\lambda)$ para varios valores de λ:

λ	$SS_E(\lambda)$
−1.00	7922.11
−0.50	687.10
−0.25	232.52
0.00	91.96
0.25	46.99
0.50	35.42
0.75	40.61
1.00	62.08
1.25	109.82
1.50	208.12

En la figura 14-1 se muestra una gráfica de los valores próximos al mínimo, en la que se observa que $\lambda \simeq 0.52$ produce un valor mínimo de aproximadamente $SS_E(\lambda) = 35.00$. Un intervalo de confianza aproximado de 95% para λ se encuentra calculando la cantidad SS^* de la ecuación 14-2 de la siguiente manera:

$$
\begin{aligned}
SS^* &= SS_E(\lambda)\left(1+\frac{t^2_{0.025,20}}{20}\right) \\
&= 35.00\left[1+\frac{(2.086)^2}{20}\right] \\
&= 42.61
\end{aligned}
$$

Al representar SS^* en la gráfica de la figura 14-1 y al leer los puntos de la escala λ donde esta recta interseca la curva, se obtienen los límites de confianza inferior y superior de λ de $\lambda^- = 0.27$ y $\lambda^+ = 0.77$. Puesto que estos límites de confianza no incluyen el valor 1, es correcto el uso de una transformación, y la transformación de la raíz cuadrada ($\lambda = 0.50$) que se usó en realidad se justifica con facilidad.

....

Algunos programas de computadora incluyen el procedimiento de Box-Cox para seleccionar una transformación de la familia de potencias. En la figura 14-2 se presenta la salida de este procedimiento como se implementa en *Design-Expert* para los datos de la descarga pico. Los resultados concuerdan en gran medida con los cálculos manuales resumidos en el ejemplo 14-1. Observe que la escala vertical de la gráfica de la figura 14-2 es $\ln[SS_E(\lambda)]$.

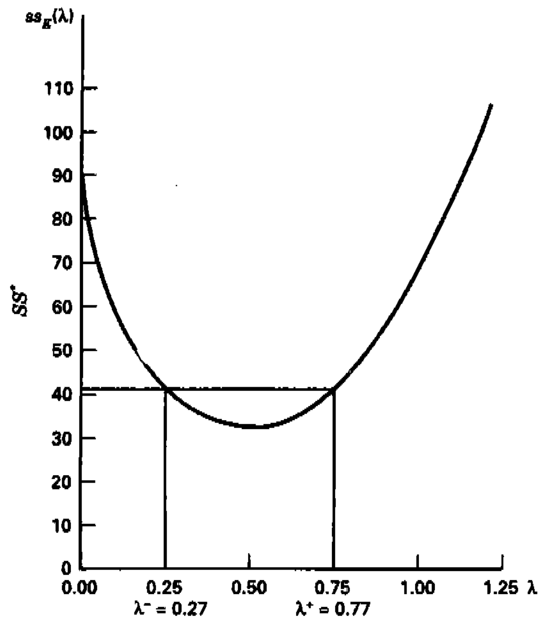


Figura 14-1 Gráfica de $SS_E(\lambda)$ contra λ para el ejemplo 14-1.

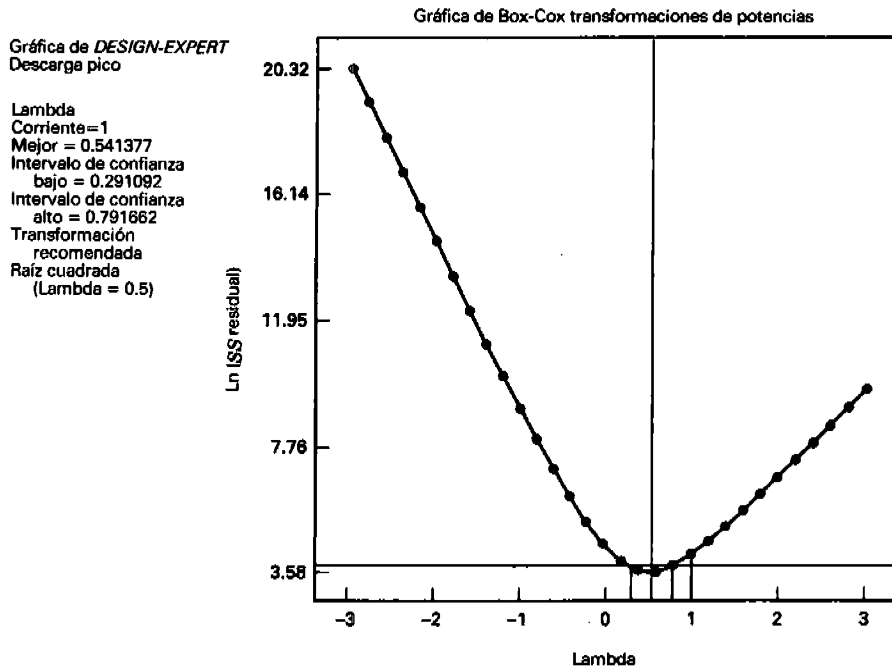


Figura 14-2 Salida de *Design-Expert* para el procedimiento de Box-Cox.

14-1.2 Modelo lineal generalizado

Con frecuencia las transformaciones de datos son una forma muy eficaz de abordar el problema de las respuestas no normales y de la desigualdad asociada de la varianza. Como se ha visto en la sección anterior, el método de Box-Cox es una forma sencilla y eficaz de seleccionar la forma de la transformación. Sin embargo, puede haber problemas asociados con el uso de una transformación de datos.

Un problema es que el experimentador puede sentirse incómodo al trabajar con la respuesta en la escala transformada. Es decir, el experimentador se interesa en el número de defectos, no en la *raíz cuadrada* del número de defectos, o en la resistividad en lugar del *logaritmo* de la resistividad. Por otra parte, si una transformación en realidad tiene éxito y mejora el análisis y el modelo asociado de la respuesta, los experimentadores adoptarán por lo general con rapidez la nueva métrica.

Un problema más serio es que una transformación puede resultar en un valor sin sentido para la variable de respuesta en alguna porción del espacio de los factores del diseño que es de interés para el experimentador. Por ejemplo, suponga que se ha usado la transformación de la raíz cuadrada en un experimento que incluye el número de defectos observados en obleas de semiconductores, y para alguna porción de la región de interés la raíz cuadrada predicha del conteo de defectos es negativa. Es probable que esto suceda en situaciones en las que el número real de defectos observados es pequeño. Por consiguiente, el modelo del experimento ha producido una predicción evidentemente no confiable justo en la región donde sería deseable que este modelo tuviera un buen desempeño predictivo.

Por último, como se señaló en la sección 14-1.1, es frecuente el uso de transformaciones a fin de estabilizar la varianza, inducir la normalidad y simplificar el modelo. No existe la seguridad de que una transformación conseguirá eficazmente todos estos objetivos al mismo tiempo.

Una alternativa del enfoque típico de la transformación de datos seguida del análisis estándar de mínimos cuadrados de la respuesta transformada es usar el **modelo lineal generalizado.** Se trata de un enfoque desarrollado por Nelder y Wedderburn [87] que en esencia unifica modelos lineales y no lineales con respuestas normales y no normales. McCullagh y Nelder [76] ofrecen un completo estudio de los modelos lineales generalizados y Myers y Montgomery [85b] proporcionan un tutorial. También se presentan detalles adicionales en el material suplementario del texto de este capítulo. Se ofrecerá un panorama general de los conceptos y se ilustrarán con dos ejemplos breves.

Un modelo lineal generalizado es básicamente un modelo de regresión (el modelo de un diseño experimental también es un modelo de regresión). Como todos los modelos de regresión, está constituido por un componente aleatorio (lo que se ha llamado generalmente el término del error) y una función de los factores del diseño (las x) y algunos parámetros desconocidos (las β). En un modelo de regresión lineal de la teoría normal estándar se escribe

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \varepsilon \tag{14-3}$$

donde se supone que el término del error ε tiene una distribución normal con media cero y varianza constante, y la media de la variable de respuesta y es

$$E(y) = \mu = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k = \mathbf{x}'\boldsymbol{\beta} \tag{14-4}$$

A la parte $\mathbf{x}'\boldsymbol{\beta}$ de la ecuación 14-4 se le llama **predictor lineal.** El modelo lineal generalizado contiene la ecuación 14-3 como un caso especial.

En un modelo lineal generalizado, la variable de respuesta puede tener cualquier distribución que sea un miembro de la **familia exponencial.** Esta familia incluye las distribuciones **normal, de Poisson, bi-**

nomial, exponencial y **gamma,** por lo que la familia exponencial es una colección rica y flexible de distribuciones aplicables en muchas situaciones experimentales. Además, la relación entre la media de la respuesta μ y el predictor lineal $\mathbf{x}'\boldsymbol{\beta}$ se determina por una **función de enlace.**

$$g(\mu)=\mathbf{x}'\boldsymbol{\beta} \tag{14-5}$$

El modelo de regresión que representa la respuesta media está dado entonces por

$$E(y)=\mu=g^{-1}(\mathbf{x}'\boldsymbol{\beta}) \tag{14-6}$$

Por ejemplo, a la función de enlace que lleva al modelo de regresión lineal ordinario en la ecuación 14-3 se le llama **enlace identidad,** ya que $\mu = g^{-1}(\mathbf{x}'\boldsymbol{\beta}) = \mathbf{x}'\boldsymbol{\beta}$. Como otro ejemplo, el **enlace log (logarítmico)**

$$\ln(\mu)=\mathbf{x}'\boldsymbol{\beta} \tag{14-7}$$

produce el modelo

$$\mu=e^{\mathbf{x}'\boldsymbol{\beta}} \tag{14-8}$$

El enlace logarítmico se usa con frecuencia con datos de conteos (respuesta de Poisson) y con respuestas continuas que presentan una distribución que tiene una cola larga a la derecha (la distribución exponencial o gamma). Otra función de enlace importante que se usa con datos binomiales es el **enlace logit**

$$\ln\left(\frac{\mu}{1-\mu}\right)=\mathbf{x}'\boldsymbol{\beta} \tag{14-9}$$

Esta elección de la función de enlace lleva al modelo

$$\mu=\frac{1}{1+e^{\mathbf{x}'\boldsymbol{\beta}}} \tag{14-10}$$

Hay muchas elecciones posibles de la función de enlace, pero debe ser siempre monótona y diferenciable. Observe asimismo que en un modelo lineal generalizado, la varianza de la variable de respuesta no tiene que ser una constante; puede ser una función de la media (y de las variables predictoras a través de la función de enlace). Por ejemplo, si la respuesta es de Poisson, la varianza de la respuesta es exactamente igual a la media.

Para usar un modelo lineal generalizado en la práctica, el experimentador debe especificar una distribución de la respuesta y una función de enlace. Después se hace el ajuste del modelo o la estimación de los parámetros por el método de máxima verosimilitud, el cual para la familia exponencial resulta ser una versión iterativa de los mínimos cuadrados ponderados. Para los modelos de regresión lineal o de diseños experimentales ordinarios con una variable de respuesta normal, esto se reduce a los mínimos cuadrados estándares. Utilizando un enfoque que es análogo al análisis de varianza de datos de la teoría normal, pueden hacerse inferencias y la verificación de diagnósticos para un modelo lineal generalizado. Referirse a Myers y Montgomery [85b] para los detalles y ejemplos. Dos paquetes de software que soportan el modelo lineal generalizado son SAS (PROC GENMOD) y S-PLUS.

EJEMPLO 14-2 ··

El experimento de los defectos en las rejillas

En el problema 8-29 se introdujo un experimento para estudiar los efectos de nueve factores sobre los defectos en rejillas de planchas moldeadas de recuadros abiertos. Bisgaard y Fuller realizaron un interesante y útil análisis de estos datos para ilustrar el valor de la transformación de datos en un experimento

Tabla 14-1 Análisis de mínimos cuadrados y del modelo lineal generalizado para el experimento rejilla de recuadros abiertos

	Utilizando métodos de mínimos cuadrados con la transformación de los datos de la raíz cuadrada modificada de Freeman y Tukey				Modelo lineal generalizado [respuesta de Poisson, enlace log (logarítmico)]		Longitud del intervalo de confianza del 95%	
	Transformados		No transformados					
Observación	Valor predicho	Intervalo de confianza de 95%	Valor predicho	Intervalo de confianza de 95%	Valor predicho	Intervalo de confianza de 95%	Mínimos cuadrados	GLM (modelo lineal generalizado)
1	5.50	(4.14, 6.85)	29.70	(16.65, 46.41)	51.26	(42.45, 61.90)	29.76	19.45
2	3.95	(2.60, 5.31)	15.12	(6.25, 27.65)	11.74	(8.14, 16.94)	21.39	8.80
3	1.52	(0.17, 2.88)	1.84	(1.69, 7.78)	1.12	(0.60, 2.08)	6.09	1.47
4	3.07	(1.71, 4.42)	8.91	(2.45, 19.04)	4.88	(2.87, 8.32)	16.59	5.45
5	1.52	(0.17, 2.88)	1.84	(1.69, 7.78)	1.12	(0.60, 2.08)	6.09	1.47
6	3.07	(1.71, 4.42)	8.91	(2.45, 19.04)	4.88	(2.87, 8.32)	16.59	5.45
7	5.50	(4.14, 6.85)	29.70	(16.65, 46.41)	51.26	(42.45, 61.90)	29.76	19.45
8	3.95	(2.60, 5.31)	15.12	(6.25, 27.65)	11.74	(8.14, 16.94)	21.39	8.80
9	1.08	(−0.28, 2.43)	0.71	(*, 5.41)	0.81	(0.42, 1.56)	*	1.13
10	−0.47	(−1.82, 0.89)	*	(*, 0.36)	0.19	(0.09, 0.38)	*	0.29
11	1.96	(0.61, 3.31)	3.36	(0.04, 10.49)	1.96	(1.16, 3.30)	10.45	2.14
12	3.50	(2.15, 4.86)	11.78	(4.13, 23.10)	8.54	(5.62, 12.98)	18.96	7.35
13	1.96	(0.61, 3.31)	3.36	(0.04, 10.49)	1.96	(1.16, 3.30)	10.45	2.14
14	3.50	(2.15, 4.86)	11.78	(4.13, 23.10)	8.54	(5.62, 12.98)	18.97	7.35
15	1.08	(−0.28, 2.43)	0.71	(*, 5.41)	0.81	(0.42, 1.56)	*	1.13
16	−0.47	(−1.82, 0.89)	*	(*, 0.36)	0.19	(0.09, 0.38)	*	0.29

diseñado. Como se señaló en el inciso *f* del problema 8-29, los autores utilizaron una modificación de la transformación de la raíz cuadrada que llevó al modelo

$$\widehat{(\sqrt{y}+\sqrt{y+1})/2} = 2.513 - 0.996x_4 - 1.21x_6 - 0.772x_2x_7$$

donde, como de costumbre, las x representan los factores del diseño codificados. Esta transformación hace un excelente trabajo para estabilizar la varianza del número de defectos. En las dos primeras secciones de la tabla 14-1 se presenta parte de la información acerca de este modelo. Bajo el encabezado "Transformados", la primera columna contiene la respuesta predicha. Observe que hay dos valores predichos negativos. El encabezado "No transformados" presenta los valores predichos no transformados, junto con los intervalos de confianza de 95% para la respuesta media en cada uno de los 16 puntos del diseño. Puesto que hubo algunos valores predichos negativos, así como límites de confianza inferiores negativos, no fue posible calcular los valores de todas las entradas de esta sección de la tabla.

La respuesta es en esencia una raíz cuadrada del conteo de los defectos. Un valor predicho negativo es claramente ilógico. Observe que esto ocurre donde los conteos observados fueron pequeños. Si es importante usar el modelo para predecir el desempeño en esta región, el modelo puede ser no confiable. Esto *no* deberá tomarse como una crítica del experimento original ni del análisis de Bisgaard y Fuller. Fue un experimento de exploración en extremo exitoso que definió con toda claridad las variables importantes del proceso. La predicción no fue una de las metas originales, y tampoco fue el objetivo del análisis realizado por Bisgaard y Fuller.

Sin embargo, si *hubiera* sido importante obtener un modelo de predicción, probablemente un modelo lineal generalizado habría sido una buena alternativa para el enfoque de la transformación. Myers y Montgomery usan un enlace log (logarítmico) (ecuación 14-7) y una respuesta de Poisson para ajustar exactamente el mismo predictor lineal dado por Bisgaard y Fuller. Esto produce el modelo

$$\hat{y} = e^{(1.128 - 0.896x_4 - 1.176x_6 - 0.737x_2x_7)}$$

La tercera sección de la tabla 14-1 contiene los valores predichos de este modelo y los intervalos de confianza de 95% para la respuesta media en cada punto del diseño (obtenida con el procedimiento PROC GENMOD de SAS). No hay valores predichos negativos (lo cual se asegura con la elección de la función de enlace) ni límites de confianza inferiores negativos. En la última sección de la tabla se comparan las longitudes de los intervalos de confianza de 95% para la respuesta no transformada y el modelo lineal generalizado (GLM). Observe que los intervalos de confianza del modelo lineal generalizado son uniformemente *más cortos* que sus contrapartes de mínimos cuadrados. Esto es un sólido indicio de que el enfoque del modelo lineal generalizado ha explicado la variabilidad y ha producido un modelo superior en comparación con el enfoque de la transformación.

EJEMPLO 14-3

El experimento del hilado de estambre

En la tabla 14-2 se presenta un diseño factorial 3^3 que se realizó para investigar el desempeño de un hilado de estambre bajo ciclos de carga repetida. El experimento se describe completo en Box y Draper [16b]. La respuesta es el número de ciclos hasta una falla. De manera típica, los datos de confiabilidad como éstos son no negativos y continuos, y con frecuencia tienen una distribución con una cola derecha alargada.

Los datos se analizaron inicialmente utilizando el enfoque estándar (mínimos cuadrados), y la transformación de datos fue necesaria para estabilizar la varianza. Se encuentra que el logaritmo de los datos

Tabla 14-2 El experimento del hilado de estambre

Corrida	x_1	x_2	x_3	Ciclos hasta una falla	Logaritmo de los ciclos hasta una falla
1	−1	−1	−1	674	2.83
2	−1	−1	0	370	2.57
3	−1	−1	1	292	2.47
4	−1	0	−1	338	253
5	−1	0	0	266	2.42
6	−1	0	1	210	2.32
7	−1	1	−1	170	2.23
8	−1	1	0	118	2.07
9	−1	1	1	90	1.95
10	0	−1	−1	1414	3.15
11	0	−1	0	1198	3.08
12	0	−1	1	634	2.8
13	0	0	−1	1022	3.01
14	0	0	0	620	2.79
15	0	0	1	438	2.64
16	0	1	−1	442	2.65
17	0	1	0	332	2.52
18	0	1	1	220	2.34
19	1	−1	−1	3636	3.56
20	1	−1	0	3184	3.5
21	1	−1	1	2000	3.3
22	1	0	−1	1568	3.19
23	1	0	0	1070	3.03
24	1	0	1	566	2.75
25	1	1	−1	1140	3.06
26	1	1	0	884	2.95
27	1	1	1	360	2.56

de los ciclos hasta una falla produce un modelo adecuado en términos del ajuste global del modelo, así como gráficas satisfactorias de los residuales. El modelo es

$$\log \hat{y} = 2.751 + 0.3617x_1 - 0.2739x_2 - 0.1711x_3$$

o en términos de la respuesta original, ciclos hasta una falla,

$$\hat{y} = 10^{2.751+0.3617x_1-0.2739x_2-0.1711x_3}.$$

Este experimento se analizó también utilizando el modelo lineal generalizado, seleccionando la distribución gamma para la respuesta y el enlace log (logarítmico). Se usó exactamente la misma forma del modelo encontrada por el análisis de mínimos cuadrados de la respuesta con la transformación logarítmica. El modelo que resultó es

$$\hat{y} = e^{6.3489+0.8425x_1-0.6313x_2-0.3851x_3}.$$

En la tabla 14-3 se presentan los valores predichos del modelo de mínimos cuadrados y del modelo lineal generalizado, junto con los intervalos de confianza de 95% para la respuesta media de cada uno de los 27

Tabla 14-3 Análisis del modelo de mínimos cuadrados y del modelo lineal generalizado para el experimento del hilado de estambre

	Métodos de mínimos cuadrados con la transformación logarítmica de los datos				Modelo lineal generalizado		Longitud del intervalo de confianza del 95%	
	Transformados		No transformados					
Observación	Valor predicho	Intervalo de confianza de 95%	Valor predicho	Intervalo de confianza de 95%	Valor predicho	Intervalo de confianza de 95%	Mínimos cuadrados	GLM (modelo lineal generalizado)
1	2.83	(2.76, 2.91)	682.50	(573.80, 811.52)	680.52	(583.83, 793.22)	237.67	209.39
2	2.66	(2.60, 2.73)	460.26	(397.01, 533.46)	463.00	(407.05, 526.64)	136.45	119.59
3	2.49	(2.42, 2.57)	310.38	(260.98, 369.06)	315.01	(271.49, 365.49)	108.09	94.00
4	2.56	(2.50, 2.62)	363.25	(313.33, 421.11)	361.96	(317.75, 412.33)	107.79	94.58
5	2.39	(2.34, 2.44)	244.96	(217.92, 275.30)	246.26	(222.55, 272.51)	57.37	49.96
6	2.22	(2.15, 2.28)	165.20	(142.50, 191.47)	167.55	(147.67, 190.10)	48.97	42.42
7	2.29	(2.21, 2.36)	193.33	(162.55, 229.93)	192.52	(165.69, 223.70)	67.38	58.01
8	2.12	(2.05, 2.18)	130.38	(112.46, 151.15)	130.98	(115.43, 148.64)	38.69	33.22
9	1.94	(1.87, 2.02)	87.92	(73.93, 104.54)	89.12	(76.87, 103.32)	30.62	26.45
10	3.20	(3.13, 3.26)	1569.28	(1353.94, 1819.28)	1580.00	(1390.00, 1797.00)	465.34	407.00
11	3.02	(2.97, 3.08)	1058.28	(941.67, 1189.60)	1075.00	(972.52, 1189.00)	247.92	216.48
12	2.85	(2.79, 2.92)	713.67	(615.60, 827.37)	731.50	(644.35, 830.44)	211.77	186.09
13	2.92	(2.87, 2.97)	835.41	(743.19, 938.86)	840.54	(759.65, 930.04)	195.67	170.39
14	2.75	(2.72, 2.78)	563.25	(523.24, 606.46)	571.87	(536.67, 609.38)	83.22	72.70
15	2.58	(2.53, 2.63)	379.84	(337.99, 426.97)	389.08	(351.64, 430.51)	88.99	78.87
16	2.65	(2.58, 2.71)	444.63	(383.53, 515.35)	447.07	(393.81, 507.54)	131.82	113.74
17	2.48	(2.43, 2.53)	299.85	(266.75, 336.98)	304.17	(275.13, 336.28)	70.23	61.15
18	2.31	(2.24, 2.37)	202.16	(174.42, 234.37)	206.95	(182.03, 235.27)	59.95	53.23
19	3.56	(3.48, 3.63)	3609.11	(3034.59, 4292.40)	3670.00	(3165.00, 4254.00)	1257.81	1089.00
20	3.39	(3.32, 3.45)	2433.88	(2099.42, 2821.63)	2497.00	(2200.00, 2833.00)	722.21	633.00
21	3.22	(3.14, 3.29)	1641.35	(1380.07, 1951.64)	1699.00	(1462.00, 1974.00)	571.57	512.00
22	3.28	(3.22, 3.35)	1920.88	(1656.91, 2226.90)	1952.00	(1720.00, 2215.00)	569.98	495.00
23	3.11	(3.06, 3.16)	1295.39	(1152.66, 1455.79)	1328.00	(1200.00, 1470.00)	303.14	270.00
24	2.94	(2.88, 3.01)	873.57	(753.53, 1012.74)	903.51	(793.15, 1029.00)	259.22	235.85
25	3.01	(2.93, 3.08)	1022.35	(859.81, 1215.91)	1038.00	(894.79, 1205.00)	356.10	310.21
26	2.84	(2.77, 2.90)	689.45	(594.70, 799.28)	706.34	(620.99, 803.43)	204.58	182.44
27	2.67	(2.59, 2.74)	464.94	(390.93, 552.97)	480.57	(412.29, 560.15)	162.04	147.86

puntos del diseño. La comparación de las longitudes de los intervalos de confianza revela que es posible que el modelo lineal generalizado sea un mejor predictor que el modelo de mínimos cuadrados.

Los modelos lineales generalizados han encontrado amplia aplicación en la investigación y el desarrollo biomédico y farmacéutico. Conforme más paquetes de software incluyan esta capacidad, encontrará una aplicación más amplia en el ámbito de la investigación y el desarrollo industrial general.

14-2 DATOS NO BALANCEADOS EN UN DISEÑO FACTORIAL

El centro de atención principal de este libro ha sido el análisis de **diseños factoriales balanceados,** es decir, los casos en que en cada celda hay el mismo número n de observaciones. Sin embargo, es común encontrar situaciones en las que el número de observaciones en las celdas son desiguales. Estos **diseños factoriales no balanceados** ocurren por varias razones. Por ejemplo, el experimentador puede haber diseñado inicialmente un experimento balanceado, pero debido a problemas imprevistos cuando se corre el experimento, los cuales resultan en la pérdida de algunas observaciones, termina trabajando con datos no balanceados. Por otra parte, algunos experimentos no balanceados se diseñan expresamente de este modo. Por ejemplo, ciertas combinaciones de tratamientos pueden ser más costosas o más difíciles de correr que otras, por lo que pueden hacerse menos observaciones en esas celdas. De manera alternativa, algunas combinaciones de tratamientos pueden ser de mayor interés para el experimentador debido a que representan condiciones nuevas o no exploradas, por lo que puede optar por obtener réplicas adicionales de dichas celdas.

La propiedad de ortogonalidad de los efectos principales y las interacciones, presente en los datos balanceados, no es válida en el caso no balanceado. Esto significa que las técnicas del análisis de varianza usual no son aplicables. Por consiguiente, el análisis de factoriales no balanceados es mucho más difícil que el de los diseños balanceados.

En esta sección se ofrece un breve panorama general de los métodos para abordar los factoriales no balanceados, centrando la atención en el caso del modelo de efectos fijos con dos factores. Suponga que el número de observaciones en la celda ij-ésima es n_{ij}. Además, sea $n_{i.} = \Sigma_{j=1}^{b}\ n_{ij}$ el número de observaciones en el renglón i-ésimo (el nivel i-ésimo del factor A), sea $n_{.j} = \Sigma_{i=1}^{a}\ n_{ij}$ el número de observaciones de la columna j-ésima (el nivel j-ésimo del factor B) y sea $n_{..} = \Sigma_{i=1}^{a}\ \Sigma_{j=1}^{b}\ n_{ij}$ el número total de observaciones.

14-2.1 Datos proporcionales: un caso sencillo

Una de las situaciones que incluye datos no balanceados presenta escasa dificultad para el análisis; se trata del caso de los **datos proporcionales.** Es decir, el número de observaciones en la celda ij-ésima es

$$n_{ij} = \frac{n_{i.}n_{.j}}{n_{..}} \tag{14-11}$$

Esta condición implica que el número de observaciones en dos renglones o columnas cualesquiera es proporcional. Cuando ocurren datos proporcionales, puede emplearse el análisis de varianza estándar. Sólo

es necesario hacer modificaciones menores en las fórmulas del cálculo manual de las sumas de cuadrados, las cuales quedan como

$$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n_{ij}} y_{ijk}^2 - \frac{y_{...}^2}{n_{..}}$$

$$SS_A = \sum_{i=1}^{a} \frac{y_{i..}^2}{n_{i.}} - \frac{y_{...}^2}{n_{..}}$$

$$SS_B = \sum_{j=1}^{b} \frac{y_{.j.}^2}{n_{.j}} - \frac{y_{...}^2}{n_{..}}$$

$$SS_{AB} = \sum_{i=1}^{a}\sum_{j=1}^{b} \frac{y_{ij.}^2}{n_{ij}} - \frac{y_{...}^2}{n_{..}} - SS_A - SS_B$$

$$SS_E = SS_T - SS_A - SS_B - SS_{AB}$$

$$= \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n_{ij}} y_{ijk}^2 - \sum_{i=1}^{a}\sum_{j=1}^{b} \frac{y_{ij.}^2}{n_{ij}}$$

Como un ejemplo de datos proporcionales, considere el experimento del diseño de la batería del ejemplo 5-1. En la tabla 14-4 se muestra una versión modificada de los datos originales. Desde luego, los datos son proporcionales; por ejemplo, en la celda 1,1 se tienen

$$n_{11} = \frac{n_{1.}n_{.1}}{n_{..}} = \frac{10(8)}{20} = 4$$

observaciones. Los resultados que se obtienen al aplicar el análisis de varianza usual a estos datos se presentan en la tabla 14-5. Tanto el tipo de material como la temperatura son significativos, lo cual concuerda con el análisis del conjunto completo de datos del ejemplo 5-1. Sin embargo, la interacción que se observó en el ejemplo 5-1 no está presente.

14-2.2 Métodos aproximados

Cuando los datos no balanceados no se apartan demasiado del caso balanceado, en ocasiones es posible usar **procedimientos aproximados** que convierten el problema no balanceado en uno balanceado. Esto

Tabla 14-4 Experimento del diseño de la batería con datos proporcionales

Tipo de material	Temperatura, °F 15	70	125	
1	$n_{11} = 4$ 130 155 74 180	$n_{12} = 4$ 34 40 80 75	$n_{13} = 2$ 70 58	$n_{1.} = 10$ $y_{1..} = 896$
2	$n_{21} = 2$ 159 126	$n_{22} = 2$ 136 115	$n_{23} = 1$ 45	$n_{2.} = 5$ $y_{2..} = 581$
3	$n_{31} = 2$ 138 160	$n_{32} = 2$ 150 139	$n_{33} = 1$ 96	$n_{3.} = 5$ $y_{3..} = 683$
	$n_{.1} = 8$ $y_{.1.} = 1122$	$n_{.2} = 8$ $y_{.2.} = 769$	$n_{.3} = 4$ $y_{.3.} = 269$	$n_{..} = 20$ $y_{...} = 2160$

Tabla 14-5 Análisis de varianza de los datos del diseño de la batería de la tabla 14-4

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0
Tipos de material	8,170.400	2	4,085.20	5.00
Temperatura	16,090.875	2	8,045.44	9.85
Interacción	5,907.725	4	1,476.93	1.18
Error	8,981.000	11	816.45	
Total	39,150.000	19		

hace, desde luego, que el análisis sea tan sólo aproximado, pero el análisis de datos balanceados es tan sencillo que con frecuencia el experimentador se ve tentado a usarlo. En la práctica, es necesario decidir cuándo los datos no son lo suficientemente diferentes del caso balanceado para hacer que el grado de aproximación introducido sea relativamente de escasa importancia. A continuación se describen brevemente algunos de estos métodos aproximados. Se supone que todas las celdas contienen al menos una observación (es decir, $n_{ij} \geq 1$).

Estimación de observaciones faltantes

Si sólo unas cuantas n_{ij} son diferentes, un procedimiento razonable es estimar los valores faltantes. Por ejemplo, considere el diseño no balanceado de la tabla 14-6. Evidentemente, estimar el único valor faltante de la celda 2,2 es un enfoque razonable. Para un modelo con interacción, la estimación del valor faltante en la celda ij-ésima que minimiza la suma de cuadrados del error es $\bar{y}_{ij.}$. Es decir, el valor faltante se estima tomando el promedio de las observaciones que están disponibles en esa celda.

El valor estimado se trata como un dato real. La única modificación del análisis de varianza es reducir los grados de libertad del error en el número de observaciones faltantes que se han estimado. Por ejemplo, si se estima el valor faltante en la celda 2,2 de la tabla 14-6, se usarían 26 grados de libertad en lugar de 27.

Apartado de datos

Considere los datos de la tabla 14-7. Observe que la celda 2,2 sólo tiene una observación más que las otras. Estimar los valores faltantes de las ocho celdas restantes quizá no sea una buena idea en este caso, ya que esto resultaría en estimaciones equivalentes a cerca de 18% de los datos finales. Una alternativa es apartar una de las observaciones de la celda 2,2, para obtener así un diseño balanceado con $n = 4$ réplicas.

La observación que se aparte deberá elegirse al azar. Además, en lugar de descartar completamente la observación, podría reintegrarse al diseño, después elegir al azar otra observación para apartarla y re-

Tabla 14-6 Los valores n_{ij} para un diseño no balanceado

	Columnas		
Renglones	1	2	3
1	4	4	4
2	4	3	4
3	4	4	4

Tabla 14-7 Los valores n_{ij} para un diseño no balanceado

	Columnas		
Renglones	1	2	3
1	4	4	4
2	4	5	4
3	4	4	4

petir el análisis. Y, se esperaría, estos dos análisis no llevarán a interpretaciones antagónicas de los datos. Si lo hacen, se sospecha que la observación que se apartó es un valor atípico o disparatado y deberá manejarse en consecuencia. En la práctica es improbable que ocurra esta confusión cuando sólo se aparta un número reducido de observaciones y la variabilidad dentro de las celdas es pequeña.

Método de las medias no ponderadas

En este enfoque, introducido por Yates [113a], los promedios de las celdas se tratan como si fueran datos y son objeto de un análisis de datos balanceados estándar para obtener las sumas de cuadrados de los renglones, las columnas y la interacción. El cuadrado medio del error se encuentra como

$$MS_E = \frac{\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n_{ij}}(y_{ijk} - \bar{y}_{ij.})^2}{n_{..} - ab} \tag{14-12}$$

Entonces, MS_E estima σ^2, la varianza de y_{ijk}, una observación individual. Sin embargo, se ha realizado un análisis de varianza de los *promedios* de las celdas, y como la varianza del promedio de la celda ij-ésima es σ^2/n_{ij}, el cuadrado medio del error que se usa en realidad en el análisis de varianza deberá ser una estimación de la varianza promedio de las $\bar{y}_{ij}$, por ejemplo

$$\bar{V}(\bar{y}_{ij.}) = \frac{\sum_{i=1}^{a}\sum_{j=1}^{b}\sigma^2/n_{ij}}{ab} = \frac{\sigma^2}{ab}\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{1}{n_{ij}} \tag{14-13}$$

Utilizando MS_E de la ecuación 14-12 para estimar σ^2 en la ecuación 14-13, se obtiene

$$MS_E' = \frac{MS_E}{ab}\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{1}{n_{ij}} \tag{14-14}$$

como el cuadrado medio del error (con $n_{..} - ab$ grados de libertad) que se usará en el análisis de varianza.

El método de las medias no ponderadas es un procedimiento aproximado porque las sumas de cuadrados de los renglones, las columnas y la interacción no se distribuyen como una variable aleatoria ji-cuadrada. La ventaja principal del método parece ser la simplicidad de los cálculos. Cuando las n_{ij} no difieren de manera radical, el método de las medias no ponderadas funciona con frecuencia razonablemente bien.

Una técnica relacionada es el **método de los cuadrados ponderados de las medias,** propuesto también por Yates [113a]. Esta técnica se basa también en las sumas de cuadrados de las medias de las celdas, pero los términos de las sumas de cuadrados se ponderan en proporción inversa a sus varianzas. Para mayores detalles de este procedimiento, ver Searle [99a] y Speed, así como Hocking y Hackney [106].

14-2.3 Método exacto

En situaciones en que los métodos aproximados no son apropiados, como cuando ocurren celdas vacías (algunas $n_{ij} = 0$) o cuando las n_{ij} presentan diferencias radicales, el experimentador debe usar un análisis exacto. El enfoque utilizado para desarrollar las sumas de cuadrados para probar los efectos principales y las interacciones consiste en representar el modelo del análisis de varianza como un **modelo de regresión,** ajustar ese modelo a los datos y usar el enfoque de la prueba general de significación de la regresión. Sin embargo, existen varias formas en que puede hacerse esto, y estos métodos pueden producir valores diferentes para las sumas de cuadrados. Además, las hipótesis que se están probando no siempre son análogos directos de las del caso balanceado, y su interpretación tampoco es siempre sencilla. Para mayor información al respecto, ver el material suplementario del texto de éste capítulo. Otras buenas referencias son Searle [99a]; Speed y Hocking [105]; Hocking y Speed [58]; Hocking, Hackney y Speed [57]; Speed, Hocking y Hackney [106]; Searle, Speed y Henderson [102]; Searle [99c]; y Milliken y Johnson [79]. El software de estadística SAS proporciona un excelente enfoque del análisis de datos no balanceados a través del procedimiento PROC GLM.

14-3 ANÁLISIS DE COVARIANZA

En los capítulos 2 y 4 se introdujo el uso del principio de la formación de bloques para mejorar la precisión con la que se hacen comparaciones entre tratamientos. La prueba t pareada fue el procedimiento ilustrado en el capítulo 2, mientras que en el capítulo 4 se presentó el diseño de bloques aleatorizados. En general, el **principio de la formación de bloques** puede usarse para eliminar el efecto de los factores perturbadores controlables. El **análisis de covarianza** es otra técnica que en ocasiones es útil para mejorar la precisión de un experimento. Suponga que en un experimento con una variable de respuesta y existe otra variable, por ejemplo x, y que y se relaciona linealmente con x. Además, suponga que x no puede ser controlada por el experimentador, pero puede observarse junto con y. A la variable x se le llama **variable concomitante** o **covariable.** El análisis de covarianza implica ajustar la variable de respuesta observada para el efecto de la variable concomitante. Si no se hace este ajuste, la variable concomitante podría inflar el cuadrado medio del error y hacer que sean más difíciles de detectar las verdaderas diferencias en la respuesta debidas a los tratamientos. Por lo tanto, el análisis de covarianza es un método de ajuste para los efectos de una variable perturbadora no controlable. Como se verá, el procedimiento es una combinación del análisis de varianza y del análisis de regresión.

Como un ejemplo de un experimento en el que puede emplearse el análisis de covarianza, considere el estudio realizado para determinar si existe una diferencia en la resistencia de una fibra de monofila-

Tabla 14-8 Datos de la resistencia a la ruptura (y = resistencia en libras y x = diámetro en 10^{-3} pulgadas)

Máquina 1		Máquina 2		Máquina 3	
y	x	y	x	y	x
36	20	40	22	35	21
41	25	48	28	37	23
39	24	39	22	42	26
42	25	45	30	34	21
49	32	44	28	32	15
207	126	216	130	180	106

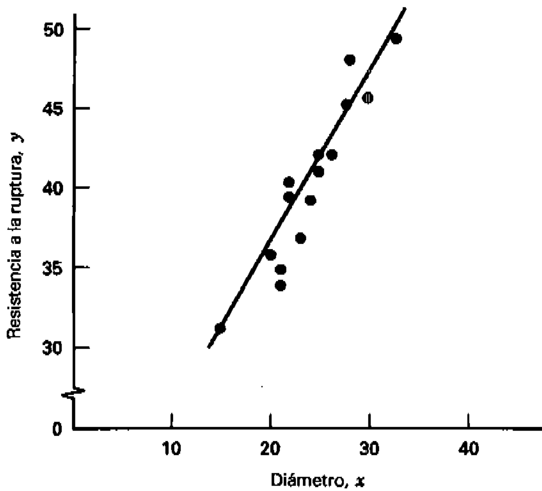


Figura 14-3 Resistencia a la ruptura (y) contra el diámetro de la fibra (x).

mento producida por tres máquinas diferentes. Los datos de este experimento se muestran en la tabla 14-8. En la figura 14-3 se presenta un diagrama de dispersión de la resistencia (y) contra el diámetro (o grosor) de la muestra. Evidentemente, la resistencia de la fibra también se afecta por su grosor; por consiguiente, una fibra más gruesa será por lo general más resistente que una delgada. El análisis de covarianza podría usarse para eliminar el efecto del grosor (x) sobre la resistencia (y) cuando se prueban las diferencias en la resistencia entre las máquinas.

14-3.1 Descripción del procedimiento

A continuación se describe el procedimiento básico para el análisis de covarianza, ilustrándolo para un experimento de un solo factor con una covariable. Suponiendo que existe una relación lineal entre la respuesta y la covariable, un modelo estadístico apropiado es

$$y_{ij} = \mu + \tau_i + \beta(x_{ij} - \bar{x}_{..}) + \varepsilon_{ij} \qquad \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, n \end{cases} \tag{14-15}$$

donde y_{ij} es la observación j-ésima de la variable de respuesta tomada bajo el tratamiento o nivel i-ésimo del único factor, x_{ij} es la medición hecha de la covariable o variable concomitante correspondiente a y_{ij} (es decir, la corrida ij-ésima), $\bar{x}_{..}$ es la media de los valores x_{ij}, μ es la media global, τ_i es el efecto del tratamiento i-ésimo, β es el coeficiente de regresión lineal que indica la dependencia de y_{ij} de x_{ij} y ε_{ij} es un componente del error aleatorio. Se supone que los errores ε_{ij} son NID(0, σ^2), que la pendiente $\beta \neq 0$ y que la verdadera relación entre y_{ij} y x_{ij} es lineal, que los coeficientes de regresión de cada tratamiento son idénticos, que la suma de los efectos de los tratamientos es cero ($\Sigma_{i=1}^{a} \tau_i = 0$) y que la variable concomitante x_{ij} no se afecta por los tratamientos.

Observe, por la ecuación 14-15, que el modelo del análisis de covarianza es una combinación de los modelos lineales empleados en el análisis de varianza y regresión. Es decir, se tienen efectos de los tratamientos $\{\tau_i\}$, como en un análisis de varianza de un solo factor, y un coeficiente de regresión β, como en una ecuación de regresión. La variable concomitante de la ecuación 14-15 se expresa ccmo $(x_{ij} - \bar{x}_{..})$

en lugar de x_{ij}, para que el parámetro μ se preserve como la media global. El modelo pudo haberse escrito como

$$y_{ij} = \mu' + \tau_i + \beta x_{ij} + \varepsilon_{ij} \qquad \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, n \end{cases} \tag{14-16}$$

donde μ' es una constante diferente de la media global, que para este modelo es $\mu' + \beta\bar{x}_{..}$. Es más común encontrar la ecuación 14-15 en la literatura sobre el tema.

Para describir el análisis, se introduce la siguiente notación:

$$S_{yy} = \sum_{i=1}^{a}\sum_{j=1}^{n} (y_{ij} - \bar{y}_{..})^2 = \sum_{i=1}^{a}\sum_{j=1}^{n} y_{ij}^2 - \frac{y_{..}^2}{an} \tag{14-17}$$

$$S_{xx} = \sum_{i=1}^{a}\sum_{j=1}^{n} (x_{ij} - \bar{x}_{..})^2 = \sum_{i=1}^{a}\sum_{j=1}^{n} x_{ij}^2 - \frac{x_{..}^2}{an} \tag{14-18}$$

$$S_{xy} = \sum_{i=1}^{a}\sum_{j=1}^{n} (x_{ij} - \bar{x}_{..})(y_{ij} - \bar{y}_{..}) = \sum_{i=1}^{a}\sum_{j=1}^{n} x_{ij} y_{ij} - \frac{(x_{..})(y_{..})}{an} \tag{14-19}$$

$$T_{yy} = n\sum_{i=1}^{a} (\bar{y}_{i.} - \bar{y}_{..})^2 = \frac{1}{n}\sum_{i=1}^{a} y_{i.}^2 - \frac{y_{..}^2}{an} \tag{14-20}$$

$$T_{xx} = n\sum_{i=1}^{a} (\bar{x}_{i.} - \bar{x}_{..})^2 = \frac{1}{n}\sum_{i=1}^{a} x_{i.}^2 - \frac{x_{..}^2}{an} \tag{14-21}$$

$$T_{xy} = n\sum_{i=1}^{a} (\bar{x}_{i.} - \bar{x}_{..})(\bar{y}_{i.} - \bar{y}_{..}) = \frac{1}{n}\sum_{i=1}^{a} (x_{i.})(y_{i.}) - \frac{(x_{..})(y_{..})}{an} \tag{14-22}$$

$$E_{yy} = \sum_{i=1}^{a}\sum_{j=1}^{n} (y_{ij} - \bar{y}_{i.})^2 = S_{yy} - T_{yy} \tag{14-23}$$

$$E_{xx} = \sum_{i=1}^{a}\sum_{j=1}^{n} (x_{ij} - \bar{x}_{i.})^2 = S_{xx} - T_{xx} \tag{14-24}$$

$$E_{xy} = \sum_{i=1}^{a}\sum_{j=1}^{n} (x_{ij} - \bar{x}_{i.})(y_{ij} - \bar{y}_{i.}) = S_{xy} - T_{xy} \tag{14-25}$$

Observe que, en general, $S = T + E$, donde los símbolos S, T y E se usan para denotar las sumas de cuadrados y los productos cruzados del total, los tratamientos y el error, respectivamente. Las sumas de cuadrados de x y y deben ser no negativas; sin embargo, las sumas de los productos cruzados (xy) pueden ser negativas.

A continuación se indica la forma en que el análisis de covarianza ajusta la variable de respuesta para el efecto de la covariable. Considere el modelo completo (ecuación 14-15). Los estimadores de mínimos cuadrados de μ, τ_i y β son $\hat{\mu} = \bar{y}_{..}$, $\hat{\tau}_i = \bar{y}_{i.} - \bar{y}_{..} - \hat{\beta}(\bar{x}_{i.} - \bar{x}_{..})$, y

$$\hat{\beta} = \frac{E_{xy}}{E_{xx}} \tag{14-26}$$

La suma de cuadrados del error en este modelo es

$$SS_E = E_{yy} - (E_{xy})^2 / E_{xx} \tag{14-27}$$

con $a(n-1) - 1$ grados de libertad. La varianza del error experimental se estima con

$$MS_E = \frac{SS_E}{a(n-1)-1}$$

Suponga ahora que no hay ningún efecto de los tratamientos. El modelo (ecuación 14-15) sería entonces

$$y_{ij} = \mu + \beta(x_{ij} - \bar{x}_{..}) + \varepsilon_{ij} \tag{14-28}$$

y puede demostrarse que los estimadores de mínimos cuadrados de μ y β son $\hat{\mu} = \bar{y}_{..}$ y $\hat{\beta} = S_{xy}/S_{xx}$. La suma de cuadrados del error en este modelo reducido es

$$SS'_E = S_{yy} - (S_{xy})^2 / S_{xx} \tag{14-29}$$

con $an - 2$ grados de libertad. En la ecuación 14-29, la cantidad $(S_{xy})^2/S_{xx}$ es la reducción de la suma de cuadrados de y obtenida a través de la regresión lineal de y sobre x. Además, observe que SS_E es menor que SS'_E [ya que el modelo (ecuación 14-15) contiene los parámetros adicionales $\{\tau_i\}$] y que la cantidad $SS'_E - SS_E$ es una reducción en la suma de cuadrados debida a las $\{\tau_i\}$. Por lo tanto, la diferencia entre SS'_E y SS_E, es decir, $SS'_E - SS_E$, proporciona una suma de cuadrados con $a - 1$ grados de libertad para probar la hipótesis de que no hay ningún efecto de los tratamientos. Por consiguiente, para probar $H_0{:}\tau_i = 0$, se calcula

$$F_0 = \frac{(SS'_E - SS_E)/(a-1)}{SS_E/[a(n-1)-1]} \tag{14-30}$$

que, si la hipótesis nula es verdadera, se distribuye como $F_{a-1, a(n-1)-1}$. Por lo tanto, $H_0{:}\tau_i = 0$ se rechaza si $F_0 > F_{\alpha, a-1, a(n-1)-1}$. También podría usarse el enfoque del valor P.

Es instructivo examinar la presentación de la tabla 14-9. En ella el análisis de covarianza se ha presentado como un análisis de varianza "ajustado". En la columna de la fuente de variación, la variabilidad total se mide por S_{yy}, con $an - 1$ grados de libertad. La fuente de variación "regresión" tiene la suma de cuadrados $(S_{xy})^2/S_{xx}$ con un grado de libertad. Si no hubiera ninguna variable concomitante, se tendría $S_{xy} = S_{xx} = E_{xy} = E_{xx} = 0$. Entonces la suma de cuadrados del error sería simplemente E_{yy} y la suma de cuadrados de los tratamientos sería $S_{yy} - E_{yy} = T_{yy}$. Sin embargo, debido a la presencia de la variable concomitante, S_{yy} y E_{yy} deben "ajustarse" para la regresión de y sobre x, como se muestra en la tabla 14-9. La suma de cuadrados del error ajustada tiene $a(n - 1) - 1$ grados de libertad en lugar de $a(n - 1)$ grados de libertad debido a que se ajusta un parámetro adicional (la pendiente β) a los datos.

Los cálculos suelen presentarse en una tabla del análisis de covarianza como la tabla 14-10. Se emplea esta presentación porque resume de manera conveniente todas las sumas de cuadrados y los productos cruzados requeridos, así como las sumas de cuadrados para probar las hipótesis acerca de los efectos de los tratamientos. Además de utilizarla para probar la hipótesis de que no hay diferencias en los efectos de los tratamientos, con frecuencia esta tabla se encuentra útil en la interpretación de los datos para presentar las medias de los tratamientos ajustadas. Estas medias ajustadas se calculan de acuerdo con

$$\bar{y}_{i.} \text{ ajustada} = \bar{y}_{i.} - \hat{\beta}(\bar{x}_{i.} - \bar{x}_{..}) \qquad i = 1, 2, \ldots, a \tag{14-31}$$

donde $\hat{\beta} = E_{xy}/E_{xx}$. Esta media de los tratamientos ajustada es el estimador de mínimos cuadrados de $\mu + \tau_i$, $i = 1, 2, \ldots, a$, en el modelo (ecuación 14-15). El error estándar de cualquier media ajustada de los tratamientos es

$$S_{\bar{y}_{i.}\text{ajustada}} = \left[MS_E \left(\frac{1}{n} + \frac{(\bar{x}_{i.} - \bar{x}_{..})^2}{E_{xx}} \right) \right]^{1/2} \tag{14-32}$$

Tabla 14-9 El análisis de covarianza como un análisis de varianza "ajustado"

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0
Regresión	$(S_{xy})^2 / S_{xx}$	1		
Tratamientos	$SS'_E - SS_E = S_{yy} - (S_{xy})^2 / S_{xx} - [E_{yy} - (E_{xy})^2 / E_{xx}]$	$a - 1$	$\frac{SS'_E - SS_E}{a-1}$	$\frac{(SS'_E - SS_E)/(a-1)}{MS_E}$
Error	$SS_E = E_{yy} - (E_{xy})^2 / E_{xx}$	$a(n-1) - 1$	$MS_E = \frac{SS_E}{a(n-1)-1}$	
Total	S_{yy}	$an - 1$		

Tabla 14-10 Análisis de covarianza de un experimento de un solo factor con una covariable

Fuente de variación	Grados de libertad	Sumas de cuadrados y productos			Ajustados para la regresión		
		x	xy	y	y	Grados de libertad	Cuadrado medio
Tratamientos	$a - 1$	T_{xx}	T_{xy}	T_{yy}			
Error	$a(n-1)$	E_{xx}	E_{xy}	E_{yy}	$SS_E = E_{yy} - (E_{xy})^2 / E_{xx}$	$a(n-1)-1$	$MS_E = \frac{SS_E}{a(n-1)-1}$
Total	$an - 1$	S_{xx}	S_{xy}	S_{yy}	$SS'_E = S_{yy} - (S_{xy})^2 / S_{xx}$	$an-2$	
Tratamientos ajustados					$SS'_E - SS_E$	$a-1$	$\frac{SS'_E - SS_E}{a-1}$

Por último, cabe recordar que se ha supuesto que el coeficiente de regresión β del modelo (ecuación 14-15) es diferente de cero. La hipótesis $H_0{:}\beta = 0$ puede probarse utilizando el estadístico de prueba

$$F_0 = \frac{(E_{xy})^2 / E_{xx}}{MS_E} \tag{14-33}$$

que bajo la hipótesis nula se distribuye como $F_{1,\, a(n-1)-1}$. Por lo tanto, $H_0{:}\beta = 0$ se rechaza si $F_0 > F_{\alpha,\, 1,\, a(n-1)-1}$.

EJEMPLO 14-4 ....

Considere el experimento descrito al principio de la sección 14-3. Tres máquinas producen una fibra de monofilamento en una fábrica textil. El ingeniero del proceso tiene interés en determinar si existe alguna diferencia en la resistencia a la ruptura de la fibra producida por las tres máquinas. Sin embargo, la resistencia de una fibra se relaciona con su diámetro, con las fibras más gruesas, siendo éstas, en general, más resistentes que las más delgadas. Se selecciona una muestra aleatoria de cinco ejemplares de prueba de fibra de cada máquina. En la tabla 14-8 se muestra la resistencia de la fibra (y) y el diámetro correspondiente (x) de cada ejemplar.

El diagrama de dispersión de la resistencia a la ruptura contra el diámetro de la fibra (figura 14-3) indica una clara tendencia a una relación lineal entre la resistencia a la ruptura y el diámetro, y parece apropiado eliminar el efecto del diámetro sobre la resistencia mediante un análisis de covarianza. Suponiendo que la relación lineal entre la resistencia a la ruptura y el diámetro es apropiada, el modelo es

$$y_{ij} = \mu + \tau_i + \beta(x_{ij} - \bar{x}_{..}) + \varepsilon_{ij} \qquad \begin{cases} i = 1, 2, 3 \\ j = 1, 2, \ldots, 5 \end{cases}$$

Utilizando las ecuaciones 14-17 a 14-25, pueden calcularse

$$S_{yy} = \sum_{i=1}^{3}\sum_{j=1}^{5} y_{ij}^2 - \frac{y_{..}^2}{an} = (36)^2 + (41)^2 + \cdots + (32)^2 - \frac{(603)^2}{(3)(5)} = 346.40$$

$$S_{xx} = \sum_{i=1}^{3}\sum_{j=1}^{5} x_{ij}^2 - \frac{x_{..}^2}{an} = (20)^2 + (25)^2 + \cdots + (15)^2 - \frac{(362)^2}{(3)(5)} = 261.73$$

$$S_{xy} = \sum_{i=1}^{3}\sum_{j=1}^{5} x_{ij} y_{ij} - \frac{(x_{..})(y_{..})}{an} = (20)(36) + (25)(41) + \cdots + (15)(32) - \frac{(362)(603)}{(3)(5)} = 282.60$$

$$T_{yy} = \frac{1}{n}\sum_{i=1}^{3} y_{i.}^2 - \frac{y_{..}^2}{an} = \frac{1}{5}[(207)^2 + (216)^2 + (180)^2] - \frac{(603)^2}{(3)(5)} = 140.40$$

$$T_{xx} = \frac{1}{n}\sum_{i=1}^{3} x_{i.}^2 - \frac{x_{..}^2}{an} = \frac{1}{5}[(126)^2 + (130)^2 + (106)^2] - \frac{(362)^2}{(3)(5)} = 66.13$$

$$T_{xy} = \frac{1}{n}\sum_{i=1}^{3} x_{i.} y_{i.} - \frac{(x_{..})(y_{..})}{an} = \frac{1}{5}[(126)(207) + (130)(216) + (106)(184)] - \frac{(362)(603)}{(3)(5)} = 96.00$$

$$E_{yy} = S_{yy} - T_{yy} = 346.40 - 140.40 = 206.00$$
$$E_{xx} = S_{xx} - T_{xx} = 261.73 - 66.13 = 195.60$$
$$E_{xy} = S_{xy} - T_{xy} = 282.60 - 96.00 = 186.60$$

Por la ecuación 14-29 se encuentra

$$\begin{aligned} SS'_E &= S_{yy} - (S_{xy})^2 / S_{xx} \\ &= 346.40 - (186.60)^2 / 261.73 \\ &= 41.27 \end{aligned}$$

con $an - 2 = (3)(5) - 2 = 13$ grados de libertad; y por la ecuación 14-27 se encuentra

$$\begin{aligned} SS_E &= E_{yy} - (E_{xy})^2 / E_{xx} \\ &= 206.00 - (186.60)^2 / 195.60 \\ &= 27.99 \end{aligned}$$

con $a(n - 1) - 1 = 3(5 - 1) - 1 = 11$ grados de libertad.

La suma de cuadrados para probar $H_0{:}\tau_1 = \tau_2 = \tau_3 = 0$ es

$$\begin{aligned} SS'_E - SS_E &= 41.27 - 27.99 \\ &= 13.28 \end{aligned}$$

con $a - 1 = 3 - 1 = 2$ grados de libertad. Estos cálculos se resumen en la tabla 14-11.

Para probar la hipótesis de que las máquinas difieren en la resistencia a la ruptura de la fibra producida, es decir, $H_0{:}\tau_i = 0$, por la ecuación 14-30 el estadístico de la prueba se calcula como

$$\begin{aligned} F_0 &= \frac{(SS'_E - SS_E)/(a-1)}{SS_E / [a(n-1)-1]} \\ &= \frac{13.28/2}{27.99/11} = \frac{6.64}{2.54} = 2.61 \end{aligned}$$

Al comparar este valor con $F_{0.10,2,11} = 2.86$, se encuentra que no puede rechazarse la hipótesis nula. El valor P de este estadístico de la prueba es $P = 0.1181$. Por lo tanto, no hay evidencia sólida de que las fibras producidas por las tres máquinas difieran en la resistencia a la ruptura.

La estimación del coeficiente de regresión se calcula con la ecuación 14-26 como

$$\hat{\beta} = \frac{E_{xy}}{E_{xx}} = \frac{186.60}{195.60} = 0.9540$$

La hipótesis $H_0{:}\beta = 0$ puede probarse usando la ecuación 14-33. El estadístico de prueba es

$$F_0 = \frac{(E_{xy})^2 / E_{xx}}{MS_E} = \frac{(186.60)^2 / 195.60}{2.54} = 70.08$$

y puesto que $F_{0.01,1,11} = 9.65$, se rechaza la hipótesis de que $\beta = 0$. Por lo tanto, existe una relación lineal entre la resistencia a la ruptura y el diámetro, y el ajuste proporcionado por el análisis de covarianza fue necesario.

Tabla 14-11 Análisis de covarianza de los datos de la resistencia a la ruptura

Fuente de variación	Grados de libertad	Sumas de cuadrados y productos			Ajustados para la región				
		x	xy	y	y	Grados de libertad	Cuadrado medio	F_0	Valor P
Máquinas	2	66.13	96.00	140.40					
Error	12	195.60	186.60	206.00	27.99	11	2.54		
Total	14	261.73	282.60	346.40	41.27	13			
Máquinas ajustadas					13.28	2	6.64	2.61	0.1181

Las medias de los tratamientos ajustadas pueden calcularse con la ecuación 14-31. Estas medias ajustadas son

$$\bar{y}_{1.} \text{ ajustada} = \bar{y}_{1.} - \hat{\beta}(\bar{x}_{1.} - \bar{x}_{..})$$
$$= 41.40 - (0.9540)(25.20 - 24.13) = 40.38$$

$$\bar{y}_{2.} \text{ ajustada} = \bar{y}_{2.} - \hat{\beta}(\bar{x}_{2.} - \bar{x}_{..})$$
$$= 43.20 - (0.9540)(26.00 - 24.13) = 41.42$$

y

$$\bar{y}_{3.} \text{ ajustada} = \bar{y}_{3.} - \hat{\beta}(\bar{x}_{3.} - \bar{x}_{..})$$
$$= 36.00 - (0.9540)(21.20 - 24.13) = 38.80$$

Al comparar las medias ajustadas con las medias no ajustadas de los tratamientos (las $\bar{y}_{i.}$), se observa que las medias ajustadas se encuentran mucho más próximas entre sí, una indicación más de que el análisis de covarianza fue necesario.

Un supuesto básico en el análisis de covarianza es que los tratamientos no influyen en la covariable x, ya que la técnica elimina el efecto de las variaciones en las $\bar{x}_{i.}$. Sin embargo, si la variabilidad en las $\bar{x}_{i.}$ se debe en parte a los tratamientos, entonces el análisis de covarianza elimina parte del efecto de los tratamientos. Por lo tanto, deberá tenerse una seguridad razonable de que los tratamientos no afectan los valores x_{ij}. En algunos experimentos esto puede ser obvio a partir de la naturaleza de la covariable, mientras que en otros puede ser más dudoso. En el ejemplo tratado aquí puede haber una diferencia en el diámetro de la fibra (x_{ij}) entre las tres máquinas. En tales casos, Cochran y Cox [26] sugieren la posible utilidad de un análisis de varianza de los valores x_{ij} para determinar la validez de este supuesto. Para el problema tratado aquí, con este procedimiento se obtiene

$$F_0 = \frac{66.13/2}{195.60/12} = \frac{33.07}{16.30} = 2.03$$

que es menor que $F_{0.10,2,12} = 2.81$, por lo que no hay razón para creer que las máquinas producen fibras con diámetros diferentes.

....

La verificación del diagnóstico del modelo de covarianza se basa en el análisis residual. Para el modelo de covarianza, los residuales son

$$e_{ij} = y_{ij} - \hat{y}_{ij}$$

donde los valores ajustados son

$$\hat{y}_{ij} = \hat{\mu} + \hat{\tau}_i + \hat{\beta}(x_{ij} - \bar{x}_{..}) = \bar{y}_{..} + [\bar{y}_{i.} - \bar{y}_{..} - \hat{\beta}(\bar{x}_{i.} - \bar{x}_{..})]$$
$$+ \hat{\beta}(x_{ij} - \bar{x}_{..}) = \bar{y}_{i.} + \hat{\beta}(x_{ij} - \bar{x}_{i.})$$

Por lo tanto,

$$e_{ij} = y_{ij} - \bar{y}_{i.} - \hat{\beta}(x_{ij} - \bar{x}_{i.}) \tag{14-34}$$

Para ilustrar el uso de la ecuación 14-34, el residual de la primera observación de la primera máquina del ejemplo 14-4 es

$$e_{11} = y_{11} - \bar{y}_{i.} - \hat{\beta}(x_{11} - \bar{x}_{1.}) = 36 - 41.4 - (0.9540)(20 - 25.2)$$
$$= 36 - 36.4392 = -0.4392$$

En la tabla siguiente se presenta una lista completa de las observaciones, los valores ajustados y los residuales:

Valor observado y_{ij}	Valor ajustado $\hat{y}_{ij}$	Residual $e_{ij} = y_{ij} - \hat{y}_{ij}$
36	36.4392	−0.4392
41	41.2092	−0.2092
39	40.2552	−1.2552
42	41.2092	0.7908
49	47.8871	1.1129
40	39.3840	0.6160
48	45.1079	2.8921
39	39.3840	−0.3840
45	47.0159	−2.0159
44	45.1079	−1.1079
35	35.8092	−0.8092
37	37.7171	−0.7171
42	40.5791	1.4209
34	35.8092	−1.8092
32	30.0852	1.9148

Los residuales se grafican contra los valores ajustados $\hat{y}_{ij}$ en la figura 14-4, contra la covariable x_{ij} en la figura 14-5 y contra las máquinas en la figura 14-6. En la figura 14-7 se muestra la gráfica de probabilidad normal de los residuales. Estas gráficas no revelan ninguna desviación importante de los supuestos, por lo que se concluye que el modelo de covarianza (ecuación 14-15) es apropiado para los datos de la resistencia a la ruptura.

Es interesante observar lo que habría ocurrido en este experimento si no se hubiera realizado el análisis de covarianza, es decir, si los datos de la resistencia a la ruptura (y) se hubieran analizado como un experimento de un solo factor en el que se ignorara la covariable x. En la tabla 14-12 se muestra el análisis de varianza de los datos de la resistencia a la ruptura. Se concluiría, con base en este análisis, que las máquinas difieren significativamente en la resistencia de la fibra producida. Es exactamente *la conclusión*

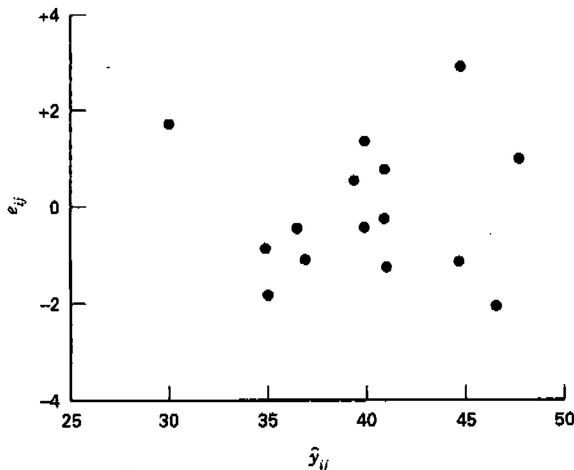


Figura 14-4 Gráfica de los residuales contra los valores ajustados del ejemplo 14-4.

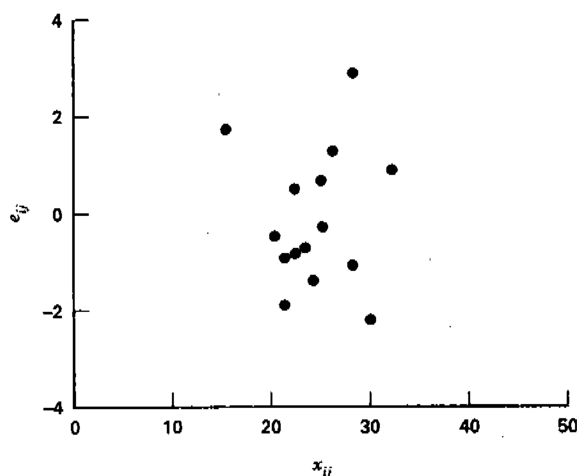


Figura 14-5 Gráfica de los residuales contra el diámetro x de la fibra en el ejemplo 14-4.

opuesta del análisis de covarianza. Si se sospechara que las máquinas difieren significativamente en su efecto sobre la resistencia de la fibra, entonces se intentaría igualar la resistencia producida por las tres máquinas. Sin embargo, en este problema las máquinas no difieren en la resistencia de la fibra producida después de que se elimina el efecto lineal del diámetro. Sería conveniente reducir la variabilidad del diámetro de la fibra dentro de las máquinas, ya que con esto probablemente se reduciría la variabilidad de la resistencia de la fibra.

14-3.2 Solución por computadora

Se cuenta con varios paquetes de software que pueden realizar el análisis de covarianza. En la tabla 14-13 se muestra la salida del procedimiento General Linear Models (modelos lineales generales) de Minitab

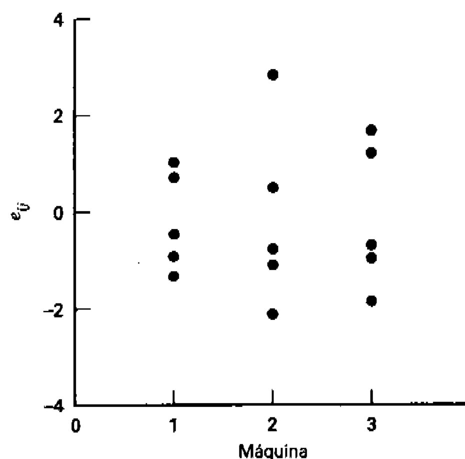


Figura 14-6 Gráfica de los residuales contra las máquinas.

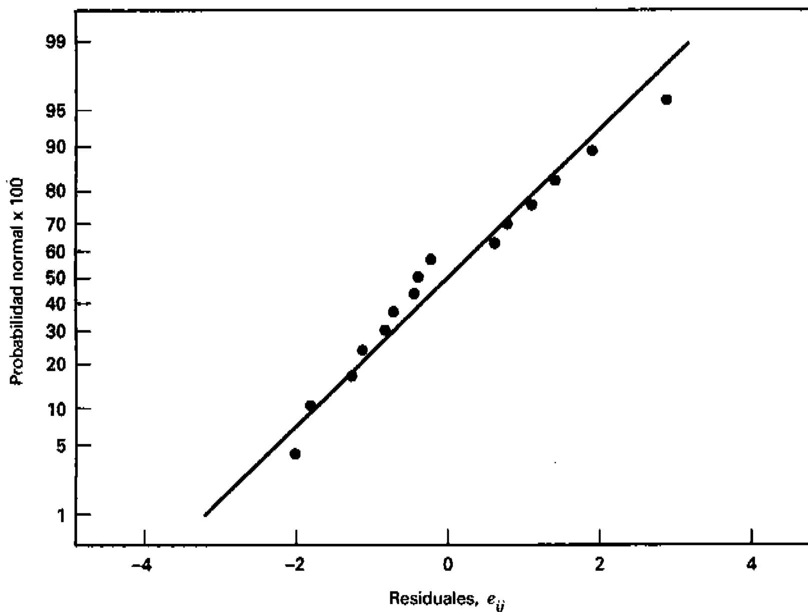


Figura 14-7 Gráfica de probabilidad normal de los residuales del ejemplo 14-4.

para los datos del ejemplo 14-4. Esta salida es muy similar a las que se presentaron anteriormente. En la sección de la salida bajo el encabezado "Análisis de varianza" ("Analysis of variance"), "SS Seq" corresponde a la partición "secuencial" de la suma de cuadrados del modelo global, es decir,

$$\begin{aligned} SS(\text{Modelo}) &= SS(\text{Diámetro}) + SS(\text{Máquina} \mid \text{Diámetro}) \\ &= 305.13 + 13.28 \\ &= 318.41 \end{aligned}$$

mientras que "SS ajustada" corresponde a la suma de cuadrados "extra" para cada factor, es decir,

$$SS(\text{Máquina} \mid \text{Diámetro}) = 13.28$$

y

$$SS(\text{Diámetro} \mid \text{Máquina}) = 178.01$$

Observe que $SS(\text{Máquina} \mid \text{Diámetro})$ es la suma de cuadrados que deberá usarse para probar que no hay ningún efecto de la máquina, y que $SS(\text{Diámetro} \mid \text{Máquina})$ es la suma de cuadrados corregida que deberá usarse para probar la hipótesis de que $\beta = 0$. Los estadísticos de la prueba de la tabla 14-13 difieren ligeramente de los que se calcularon manualmente debido al redondeo.

Tabla 14-12 Análisis incorrecto de los datos de la resistencia a la ruptura como un experimento de un solo factor

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0	Valor P
Máquinas	140.40	2	70.20	4.09	0.0442
Error	206.00	12	17.17		
Total	346.40	14			

Tabla 14-13 Salida de Minitab (análisis de covarianza) del ejemplo 14-4

Modelo lineal general

```
Factor      Type  Levels Values
Machine     fixed      3 1 2 3

Analysis of Variance for Strength, using Adjusted SS for Tests

Source      DF     Seq SS      Adj SS      Adj MS       F      P
Diameter     1     305.13      178.01      178.01   69.97  0.000
Machine      2      13.28       13.28        6.64    2.61  0.118
Error       11      27.99       27.99        2.54
Total       14     346.40

Term           Coef      StDev          T       P
Constant     17.177      2.783       6.17   0.000
Diameter     0.9540     0.1140       8.36   0.000
Machine
1            0.1824     0.5950       0.31   0.765
2            1.2192     0.6201       1.97   0.075

Means for Covariates

Covariate       Mean     StDev
Diameter       24.13     4.324

Least Squares Means for Strength

Machine        Mean     StDev
1             40.38    0.7236
2             41.42    0.7444
3             38.80    0.7879
```

El programa calcula también las medias de los tratamientos ajustadas con la ecuación 14-31 (Minitab hace referencia a éstas como las medias de mínimos cuadrados en la salida muestral) y los errores estándar. El programa comparará asimismo todos los pares de medias de tratamientos utilizando los procedimientos de comparación múltiple por pares estudiados en el capítulo 3.

14-3.3 Desarrollo mediante la prueba general de significación de la regresión

Es posible desarrollar formalmente el procedimiento para probar $H_0{:}\tau_i = 0$ en el modelo de covarianza

$$y_{ij} = \mu + \tau_i + \beta(x_{ij} - \bar{x}_{..}) + \varepsilon_{ij} \qquad \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, n \end{cases} \tag{14-35}$$

utilizando la prueba general de significación de la regresión. Considere la estimación de los parámetros del modelo (ecuación 14-15) por mínimos cuadrados. La función de mínimos cuadrados es

$$L = \sum_{i=1}^{a}\sum_{j=1}^{n} [y_{ij} - \mu - \tau_i - \beta(x_{ij} - \bar{x}_{..})]^2 \tag{14-36}$$

y a partir de $\partial L/\partial\mu = \partial L/\partial\tau_i = \partial L/\partial\beta = 0$, se obtienen las ecuaciones normales

$$\mu: an\hat{\mu} + n\sum_{i=1}^{a} \hat{\tau}_i = y_{..} \quad (14\text{-}37a)$$

$$\tau_i: n\hat{\mu} + n\hat{\tau}_i + \hat{\beta}\sum_{j=1}^{n} (x_{ij} - \bar{x}_{..}) = y_{i.} \qquad i = 1, 2, \ldots, a \quad (14\text{-}37b)$$

$$\beta: \sum_{i=1}^{a}\hat{\tau}_i \sum_{j=1}^{n} (x_{ij} - \bar{x}_{..}) + \hat{\beta}S_{xx} = S_{xy} \quad (14\text{-}37c)$$

Al sumar las a ecuaciones de la ecuación 14-37b, se obtiene la ecuación 14-37a porque $\Sigma_{i=1}^{a}\Sigma_{j=1}^{n}(x_{ij} - \bar{x}_{..}) = 0$, por lo que existe una dependencia lineal en las ecuaciones normales. Por lo tanto, es necesario aumentar las ecuaciones 14-37 con una ecuación linealmente independiente para obtener una solución. Una condición lógica es $\Sigma_{i=1}^{a}\hat{\tau}_i = 0$.

Al utilizar esta condición, de la ecuación 14-37a se obtiene

$$\hat{\mu} = \bar{y}_{..} \quad (14\text{-}38a)$$

y de la ecuación 14-37b

$$\hat{\tau}_i = \bar{y}_{i.} - \bar{y}_{..} - \hat{\beta}(\bar{x}_{i.} - \bar{x}_{..}) \quad (14\text{-}38b)$$

La ecuación 14-37c puede reescribirse como

$$\sum_{i=1}^{a} (\bar{y}_{i.} - \bar{y}_{..})\sum_{j=1}^{n} (x_{ij} - \bar{x}_{..}) - \hat{\beta}\sum_{i=1}^{a} (\bar{x}_{i.} - \bar{x}_{..})\sum_{j=1}^{n} (x_{ij} - \bar{x}_{..}) + \hat{\beta}S_{xx} = S_{xy}$$

después de sustituir para $\hat{\tau}_i$. Pero se observa que

$$\sum_{i=1}^{a} (\bar{y}_{i.} - \bar{y}_{..})\sum_{j=1}^{n} (x_{ij} - \bar{x}_{..}) = T_{xy}$$

y

$$\sum_{i=1}^{a} (\bar{x}_{i.} - \bar{x}_{..})\sum_{j=1}^{n} (x_{ij} - \bar{x}_{..}) = T_{xx}$$

Por lo tanto, la solución de la ecuación 14-37c es

$$\hat{\beta} = \frac{S_{xy} - T_{xy}}{S_{xx} - T_{xx}} = \frac{E_{xy}}{E_{xx}}$$

que fue el resultado dado anteriormente en la sección 14-3.1 en la ecuación 14-26.

La reducción en la suma de cuadrados total debida al ajuste del modelo (ecuación 14-15) puede expresarse como

$$\begin{aligned} R(\mu, \tau, \beta) &= \hat{\mu}y_{..} + \sum_{i=1}^{a} \hat{\tau}_i y_{i.} + \hat{\beta}S_{xy} \\ &= (\bar{y}_{..})y_{..} + \sum_{i=1}^{a} [\bar{y}_{i.} - \bar{y}_{..} - (E_{xy}/E_{xx})(\bar{x}_{i.} - \bar{x}_{..})]y_{i.} + (E_{xy}/E_{xx})S_{xy} \\ &= y_{..}^2/an + \sum_{i=1}^{a} (\bar{y}_{i.} - \bar{y}_{..})y_{i.} - (E_{xy}/E_{xx})\sum_{i=1}^{a} (\bar{x}_{i.} - \bar{x}_{..})y_{i.} + (E_{xy}/E_{xx})S_{xy} \\ &= y_{..}^2/an + T_{yy} - (E_{xy}/E_{xx})(T_{xy} - S_{xy}) \\ &= y_{..}^2/an + T_{yy} + (E_{xy})^2/E_{xx} \end{aligned}$$

Esta suma de cuadrados tiene $a + 1$ grados de libertad porque el rango de las ecuaciones normales es $a + 1$. La suma de cuadrados del error de este modelo es

$$\begin{aligned} SS_E &= \sum_{i=1}^{a}\sum_{j=1}^{n} y_{ij}^2 - R(\mu, \tau, \beta) \\ &= \sum_{i=1}^{a}\sum_{j=1}^{n} y_{ij}^2 - y_{..}^2 / an - T_{yy} - (E_{xy})^2 / E_{xx} \\ &= S_{yy} - T_{yy} - (E_{xy})^2 / E_{xx} \\ &= E_{yy} - (E_{xy})^2 / E_{xx} \end{aligned} \tag{14-39}$$

con $an - (a + 1) = a(n - 1) - 1$ grados de libertad. Esta cantidad se obtuvo anteriormente como la ecuación 14-27.

Considere ahora el modelo restringido a la hipótesis nula, es decir, a $H_0{:}\tau_1 = \tau_2 = \ldots = \tau_a = 0$. Este modelo reducido es

$$y_{ij} = \mu + \beta(x_{ij} - \bar{x}_{..}) + \varepsilon_{ij} \qquad \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, n \end{cases} \tag{14-40}$$

Se trata de un modelo de regresión lineal simple, y las ecuaciones normales de mínimos cuadrados para este modelo son

$$an\hat{\mu} = y_{..} \tag{14-41a}$$

$$\hat{\beta}S_{xx} = S_{xy} \tag{14-41b}$$

Las soluciones de estas ecuaciones son $\hat{\mu} = \bar{y}_{..}$ y $\hat{\beta} = S_{xy}/S_{xx}$; y la reducción en la suma de cuadrados total debida al ajuste del modelo reducido es

$$\begin{aligned} R(\mu, \beta) &= \hat{\mu}y_{..} + \hat{\beta}S_{xy} \\ &= (\bar{y}_{..})y_{..} + (S_{xy} / S_{xx})S_{xy} \\ &= y_{..}^2 / an + (S_{xy})^2 / S_{xx} \end{aligned} \tag{14-42}$$

Esta suma de cuadrados tiene dos grados de libertad.

La suma de cuadrados apropiada para probar $H_0{:}\tau_1 = \tau_2 = \ldots = \tau_a = 0$ puede encontrarse como

$$\begin{aligned} R(\tau|\mu, \beta) &= R(\mu, \tau, \beta) - R(\mu, \beta) \\ &= y_{..}^2 / an + T_{yy} + (E_{xy})^2 / E_{xx} - y_{..}^2 / an - (S_{xy})^2 / S_{xx} \\ &= S_{yy} - (S_{xy})^2 / S_{xx} - [E_{yy} - (E_{xy})^2 / E_{xx}] \end{aligned} \tag{14-43}$$

utilizando $T_{yy} = S_{yy} - E_{yy}$. Observe que $R(\tau|\mu,\beta)$ tiene $a + 1 - 2 = a - 1$ grados de libertad y que es idéntica a la suma de cuadrados dada por $SS_E' - SS_E$ en la sección 14-3.1. Por lo tanto, el estadístico de prueba para $H_0{:}\tau_i = 0$ es

$$F_0 = \frac{R(\tau|\mu, \beta)/(a-1)}{SS_E / [a(n-1)-1]} = \frac{(SS_E' - SS_E)/(a-1)}{SS_E / [a(n-1)-1]} \tag{14-44}$$

expresión que se dio anteriormente como la ecuación 14-30. Por lo tanto, utilizando la prueba general de significación de la regresión, se ha justificado el desarrollo heurístico del análisis de covarianza de la sección 14-3.1.

14-3.4 Experimentos factoriales con covariables

El análisis de covarianza puede aplicarse a estructuras de tratamientos más complejas, como los diseños factoriales. Siempre que existan datos suficientes para cada combinación de tratamientos, prácticamente cualquier estructura de tratamientos compleja puede analizarse mediante el enfoque del análisis de covarianza. A continuación se indica cómo podría usarse el análisis de covarianza en la familia más común de diseños factoriales utilizados en la experimentación industrial, los factoriales 2^k.

Al establecer el supuesto de que la covariable afecta a la variable de respuesta de manera idéntica para todas las combinaciones de tratamientos, podría construirse una tabla del análisis de covarianza similar al procedimiento dado en la sección 14-3.1. La única diferencia sería la suma de cuadrados de los tratamientos. Para un factorial 2^2 con n réplicas, la suma de cuadrados de los tratamientos (T_{yy}) sería $(1/n)$ $\Sigma_{i=1}^{2} \Sigma_{j=1}^{2} y_{ij.}^2 - y_{...}^2/(2)(2)n$. Esta cantidad es la suma de las sumas de cuadrados de los factores A, B y la interacción AB. Entonces podría hacerse la partición de la suma de cuadrados ajustada de los tratamientos en componentes de los efectos individuales, es decir, la suma de cuadrados de los efectos principales ajustados SS_A y SS_B, y una suma de cuadrados de la interacción, SS_{AB}.

El número de réplicas es un aspecto clave cuando se amplía la estructura de los tratamientos del diseño. Considere un arreglo factorial 2^3. Se necesita un mínimo de dos réplicas para evaluar todas las combinaciones de tratamientos con una covariable separada para cada combinación de tratamientos (una covariable por interacción de tratamientos). Esto es equivalente a ajustar un modelo de regresión simple a cada combinación de tratamientos o celda del diseño. Con dos observaciones por celda, un grado de libertad se usa para estimar la ordenada al origen (el efecto del tratamiento), y el otro se usa para estimar la pendiente (el efecto de la covariable). Con este modelo saturado, no se cuenta con ningún grado de libertad para estimar el error. Por lo tanto, se necesitan al menos tres réplicas para un análisis de covarianza completo, suponiendo el caso más general. Este problema se agudiza cuando se incrementa el número de celdas distintas del diseño (combinaciones de tratamientos) y las covariables.

Si el número de réplicas está limitado, pueden hacerse varios supuestos para permitir un análisis útil. El supuesto más simple (y típicamente el peor) que puede hacerse es que la covariable no tiene ningún efecto. Si la covariable, incorrectamente, deja de tomarse en consideración, el análisis completo y las conclusiones subsecuentes podrían tener graves errores. Otra elección es suponer que no hay ningún tratamiento por interacción de la covariable. Aun cuando este supuesto sea incorrecto, el efecto promedio de la covariable en todos los tratamientos seguirá incrementando la precisión de la estimación y la prueba de los efectos de los tratamientos. Una desventaja de este supuesto es que si varios niveles de los tratamientos interactúan con la covariable, los diferentes términos pueden cancelarse entre sí y el término de la covariable, si se estima solo sin ninguna interacción, puede resultar no significativo. Una tercera elección sería suponer que algunos de los factores (como algunas interacciones de dos factores y de órdenes superiores) no son significativos. Esto permite usar parte de los grados de libertad para estimar el error. Sin embargo, este curso de acción deberá emprenderse con cuidado, y los modelos subsecuentes deberán evaluarse a profundidad, ya que la estimación del error será relativamente imprecisa a menos que se le asignen suficientes grados de libertad. Con dos réplicas, cada uno de estos supuestos liberará algunos grados de libertad para estimar el error y permitirá realizar pruebas de hipótesis útiles. El supuesto que se establecerá deberá ser dictado por la situación experimental y por el riesgo que el experimentador esté dispuesto a correr. Cabe hacer notar que en la estrategia de construcción del modelo de los efectos, si se elimina el factor de uno de los tratamientos, entonces las dos "réplicas" resultantes de cada factorial 2^3 original no son en realidad réplicas. Estas "réplicas ocultas" liberan grados de libertad para la estimación de parámetros, pero no deberán usarse como réplicas para estimar el error puro porque la ejecución del diseño original quizá no se haya aleatorizado para ello.

Tabla 14-14 Datos de la respuesta y la covariable para un diseño 2^3 con 2 réplicas

A	B	C	x	y
–1	–1	–1	4.05	–30.73
1	–1	–1	0.36	9.07
–1	1	–1	5.03	39.72
1	1	–1	1.96	16.30
–1	–1	1	5.38	–26.39
1	–1	1	8.63	54.58
–1	1	1	4.10	44.54
1	1	1	11.44	66.20
–1	–1	–1	3.58	–26.46
1	–1	–1	1.06	10.94
–1	1	–1	15.53	103.01
1	1	–1	2.92	20.44
–1	–1	1	2.48	–8.94
1	–1	1	13.64	73.72
–1	1	1	–0.67	15.89
1	1	1	5.13	38.57

Para ilustrar algunas de estas ideas, considere el diseño factorial 2^3 con dos réplicas y una covariable que se muestra en la tabla 14-14. Si la variable de respuesta y se analiza sin tomar en cuenta la covariable, resulta el siguiente modelo:

$$\hat{y} = 25.03 + 11.20A + 18.05B + 7.24C - 18.91AB + 14.80AC$$

El modelo global es significativo en el nivel de $\alpha = 0.01$ con $R^2 = 0.786$ y $MS_E = 470.82$. El análisis residual no indica problemas con este modelo, excepto porque la observación con $y = 103.01$ es inusual.

Si se elige el segundo supuesto, que las pendientes son comunes con ningún tratamiento por interacción de la covariable, pueden estimarse el modelo de los efectos completo y el efecto de la covariable. La salida de Minitab (de la rutina General Linear Models) se muestra en la tabla 14-15. Observe que MS_E se ha reducido considerablemente al tomar en consideración la covariable. El análisis final resultante después de eliminar de manera secuencial cada interacción no significativa y el efecto principal C se muestra en la tabla 14-16. Este modelo reducido proporciona un MS_E todavía menor que el modelo completo con la covariable de la tabla 14-15.

Por último, podría considerarse un tercer curso de acción, suponiendo que ciertos términos de interacción son insignificantes. Se considera el modelo completo que permite pendientes diferentes entre los tratamientos y la interacción tratamiento por covariable. Se supone que no son significativas las interacciones de tres factores (tanto ABC como $ABCx$) y se usan los grados de libertad asociados con ellas para estimar el error en el modelo de los efectos más general que pueda ajustarse. Éste es con frecuencia un supuesto práctico. Las interacciones de tres factores son por lo general insignificantes en la mayoría de los ambientes experimentales. La versión actual de Minitab no puede modelar covariables que interactúan con los tratamientos, por lo que se usa PROC GLM de SAS. Las sumas de cuadrados tipo III son las sumas de cuadrados ajustadas que se necesitan. En la tabla 14-17 se presentan los resultados de SAS para este modelo.

Con un modelo casi saturado, la estimación del error será bastante imprecisa. Incluso cuando unos cuantos términos son individualmente significativos en el nivel $\alpha = 0.05$, el sentido general es que este modelo es mejor que los dos escenarios previos (basados en R^2 y el cuadrado medio del error). Debido a que el aspecto de los efectos de los tratamientos del modelo es de mayor interés, se eliminan de manera secuencial términos de la porción de la covariable del modelo a fin de agregar grados de libertad para esti-

Tabla 14-15 Análisis de covarianza de Minitab para el experimento de la tabla 14-14, suponiendo una pendiente común

Modelo lineal general

```
Factor      Type Levels Values
A          fixed      2 -1   1
B          fixed      2 -1   1
C          fixed      2 -1   1

Analysis of Variance for y, using Adjusted SS for Tests

Source      DF      Seq SS      Adj SS      Adj MS        F       P
x            1     12155.9      2521.6      2521.6    28.10   0.001
A            1      1320.7      1403.8      1403.8    15.64   0.005
B            1      3997.6      4066.2      4066.2    45.31   0.000
C            1        52.7        82.3        82.3     0.92   0.370
A*B          1      3788.3      3641.0      3641.0    40.58   0.000
A*C          1        10.2         1.1         1.1     0.01   0.913
B*C          1         5.2         8.4         8.4     0.09   0.769
A*B*C        1        33.2        33.2        33.2     0.37   0.562
Error        7       628.1       628.1        89.7
Total       15     21992.0

Term           Coef      StDev        T      P
Constant     -1.016      5.454    -0.19  0.858
x            4.9245     0.9290     5.30  0.001
```

mar el error. Si se elimina secuencialmente el término ACx seguido de BCx, el MS_E decrece a 0.7336 y varios términos no son significativos. En la tabla 14-18 se muestra el modelo final después de eliminar secuencialmente Cx, AC y BC.

Este ejemplo destaca la necesidad de contar con grados de libertad para estimar el error experimental a fin de incrementar la precisión de las pruebas de hipótesis asociadas con los términos individuales del modelo. Este proceso deberá hacerse de manera secuencial para evitar la eliminación de términos significativos enmascarados por una estimación pobre del error.

Tabla 14-16 Análisis de covarianza de Minitab, modelo reducido para el experimento de la tabla 14-14

Modelo lineal general

```
Factor      Type Levels Values
A          fixed      2 -1   1
B          fixed      2 -1   1

Analysis of Variance for y, using Adjusted SS for Tests

Source      DF      Seq SS      Adj SS      Adj MS        F       P
x            1     12155.9      8287.9      8287.9   119.43   0.000
A            1      1320.7      1404.7      1404.7    20.24   0.001
B            1      3997.6      4097.7      4097.7    59.05   0.000
A*B          1      3754.5      3754.5      3754.5    54.10   0.000
Error       11       763.3       763.3        69.4
Total       15     21992.0

Term           Coef      StDev        T      P
Constant     -1.878      3.225    -0.58  0.572
x            5.0876     0.4655    10.93  0.000
```

Tabla 14-17 Salida de PROC GLM (análisis de covarianza) de SAS para el experimento de la tabla 14-14

Dependent Variable: Y

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	13	21989.20828	1691.47756	1206.45	0.0008
Error	2	2.80406	1.40203		
Corrected Total	15	21992.01234			

R-Square	C.V.	Root MSE	Y Mean
0.999872	4.730820	1.184074	25.02895

Source	DF	Type III SS	Mean Square	F Value	Pr > F
A	1	4.6599694	4.6599694	3.32	0.2099
B	1	13.0525319	13.0525319	9.31	0.0927
C	1	35.0087994	35.0087994	24.97	0.0378
AB	1	17.1013635	17.1013635	12.20	0.0731
AC	1	0.0277472	0.0277472	0.02	0.9010
BC	1	0.4437474	0.4437474	0.32	0.6304
X	1	49.2741287	49.2741287	35.14	0.0273
AX	1	33.9024288	33.9024288	24.18	0.0390
BX	1	95.7747490	95.7747490	68.31	0.0143
CX	1	0.1283784	0.1283784	0.09	0.7908
ABX	1	336.9732676	336.9732676	240.35	0.0041
ACX	1	0.0020997	0.0020997	0.00	0.9726
BCX	1	0.0672386	0.0672386	0.05	0.8470

Tabla 14-18 Salida de PROC GLM de SAS para el experimento de la tabla 14-14, modelo reducido

Dependent Variable: Y

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	8	21986.33674	2748.29209	3389.61	0.0001
Error	7	5.67560	0.81080		
Corrected Total	15	21992.01234			

R-Square	C.V.	Root MSE	Y Mean
0.999742	3.597611	0.900444	25.02895

Source	DF	Type III SS	Mean Square	F Value	Pr > F
A	1	19.1597158	19.1597158	23.63	0.0018
B	1	38.0317496	38.0317496	46.91	0.0002
C	1	232.2435668	232.2435668	286.44	0.0001
AB	1	31.7635098	31.7635098	39.18	0.0004
X	1	240.8726525	240.8726525	297.08	0.0001
AX	1	233.3934567	233.3934567	287.86	0.0001
BX	1	550.1530561	550.1530561	678.53	0.0001
ABX	1	542.3268940	542.3268940	668.88	0.0001

Parameter	Estimate	T for H0: Parameter=0	Pr > |T|	Std Error of Estimate
Intercept	10.2438830	18.74	0.0001	0.54659908
A	2.7850330	4.86	0.0018	0.57291820
B	3.6596279	6.85	0.0002	0.53434356
C	5.4560862	16.92	0.0001	0.32237858
AB	-3.3636850	-6.26	0.0004	0.53741264
X	2.0471937	17.24	0.0001	0.11877417
AX	2.0632049	16.97	0.0001	0.12160595
BX	3.0340997	26.05	0.0001	0.11647826
ABX	-3.0342229	-25.86	0.0001	0.11732045

Al revisar los resultados obtenidos de los tres enfoques, se observa que cada método mejora de manera sucesiva el ajuste del modelo en este ejemplo. Si hay una razón fundada para creer que la covariable no interactúa con los factores, quizá sea mejor establecer ese supuesto desde el principio del análisis. Esta opción también puede ser dictada por el software. Aun cuando los paquetes de software para diseños experimentales quizá sólo tengan capacidad para modelar covariables que no interactúan con los tratamientos, el analista puede tener una oportunidad razonable de identificar los factores principales que influyen en el proceso, incluso si hay alguna covariable por interacción de tratamientos. Se observa asimismo que las pruebas usuales de la adecuación del modelo siguen siendo apropiadas y se recomiendan enérgicamente como parte del proceso de construcción del modelo del análisis de covarianza ANCOVA.

14-4 MEDICIONES REPETIDAS

En el trabajo experimental de las ciencias sociales y el comportamiento, así como en algunos aspectos de la ingeniería y las ciencias físicas, las unidades experimentales son con frecuencia personas. Debido a las discrepancias en experiencia, capacitación o formación, en algunas situaciones experimentales las diferencias en las respuestas de distintas personas al mismo tratamiento pueden ser muy grandes. A menos que esté controlada, esta variabilidad entre las personas se convertirá en parte del error experimental, y en algunos casos inflará significativamente el cuadrado medio del error, haciendo más difícil detectar las diferencias reales entre los tratamientos.

Es posible controlar esta variabilidad entre las personas utilizando un diseño en el que cada uno de los a tratamientos se usa en cada persona (o "sujeto"). A éste se le llama **diseño de mediciones repetidas.** En esta sección se ofrece una breve introducción a los experimentos de mediciones repetidas con un solo factor.

Suponga que un experimento incluye a tratamientos y que cada tratamiento se va a usar exactamente una sola vez en cada uno de n sujetos. Los datos aparecerían como en la tabla 14-19. Vea que la observación y_{ij} representa la respuesta del sujeto j al tratamiento i y que sólo se usan n sujetos. El modelo que se utiliza para este diseño es

$$y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij} \tag{14-45}$$

donde τ_i es el efecto del tratamiento i-ésimo y β_j es un parámetro asociado con el sujeto j-ésimo. Se supone que los tratamientos son fijos (de donde $\Sigma_{i=1}^{a}\tau_i = 0$) y que los sujetos empleados son una muestra aleatoria de alguna población más grande de individuos potenciales. Por lo tanto, colectivamente los sujetos representan un efecto aleatorio, por lo que se supone que la media de β_j es cero y que la varianza de β_j es σ_β^2. Puesto que el término β_j es común a todas las a mediciones del mismo sujeto, la covarianza entre y_{ij} y $y_{i'j}$ no

Tabla 14-19 Datos de un diseño de mediciones repetidas con un solo factor

	Sujeto				Totales de
Tratamiento	1	2	⋯	n	los tratamientos
1	y_{11}	y_{12}	⋯	y_{1n}	$y_{1.}$
2	y_{21}	y_{22}	⋯	y_{2n}	$y_{2.}$
⋮	⋮	⋮		⋮	⋮
a	y_{a1}	y_{a2}		y_{an}	$y_{a.}$
Totales de los sujetos	$y_{.1}$	$y_{.2}$	⋯	$y_{.n}$	$y_{..}$

es, en general, cero. Se acostumbra suponer que la covarianza entre y_{ij} y $y_{i'j}$ es constante a lo largo de todos los tratamientos y los sujetos.

Considere una partición en el análisis de varianza de la suma de cuadrados total, por ejemplo

$$\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij}-\bar{y}_{..})^2 = a\sum_{j=1}^{n}(\bar{y}_{.j}-\bar{y}_{..})^2+\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij}-\bar{y}_{.j})^2 \qquad (14\text{-}46)$$

El primer término del miembro derecho de la ecuación 14-46 puede considerarse como una suma de cuadrados que resulta de las diferencias *entre los sujetos*, y el segundo término es una suma de cuadrados de las diferencias *dentro de los sujetos*. Es decir,

$$SS_T = SS_{\text{Entre los sujetos}} + SS_{\text{Dentro de los sujetos}}$$

Las sumas de cuadrados $SS_{\text{Entre los sujetos}}$ y $SS_{\text{Dentro de los sujetos}}$ son estadísticamente independientes, con grados de libertad

$$an-1=(n-1)+n(a-1)$$

Las diferencias dentro de los sujetos dependen tanto de las diferencias en los efectos de los tratamientos como de la variabilidad no controlada (ruido o error). Por lo tanto, la suma de cuadrados resultante de las diferencias dentro de los sujetos puede descomponerse de la siguiente manera:

$$\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij}-\bar{y}_{.j})^2 = n\sum_{i=1}^{a}(\bar{y}_{i.}-\bar{y}_{..})^2+\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij}-\bar{y}_{i.}-\bar{y}_{.j}+\bar{y}_{..})^2 \qquad (14\text{-}47)$$

El primer término del miembro derecho de la ecuación 14-47 mide la contribución de la diferencia entre las medias de los tratamientos a $SS_{\text{Dentro de los sujetos}}$, y el segundo término es la variación residual debida al error. Ambos componentes de $SS_{\text{Dentro de los sujetos}}$ son independientes. Por lo tanto,

$$SS_{\text{Dentro de los sujetos}} = SS_{\text{Tratamientos}} + SS_E$$

con los grados de libertad dados por

$$n(a-1)=(a-1)+(a-1)(n-1)$$

respectivamente.

Para probar la hipótesis de que no hay ningún efecto de los tratamientos, es decir,

$$H_0: \tau_1=\tau_2=\cdots=\tau_a=0$$
$$H_1: \text{Al menos una } \tau_i \neq 0$$

se usaría el cociente

$$F_0=\frac{SS_{\text{Tratamientos}}/(a-1)}{SS_E/(a-1)(n-1)}=\frac{MS_{\text{Tratamientos}}}{MS_E} \qquad (14\text{-}48)$$

Si los errores del modelo siguen una distribución normal, entonces bajo la hipótesis nula, H_0: $\tau_i = 0$, el estadístico F_0 sigue una distribución $F_{a-1,(a-1)(n-1)}$. La hipótesis nula se rechazaría si $F_0 > F_{\alpha,a-1,(a-1)(n-1)}$.

En la tabla 14-20 se resume el procedimiento del análisis de varianza, donde se presentan también fórmulas convenientes de cálculo para las sumas de cuadrados. El lector deberá identificar el análisis de varianza de un diseño de un solo factor con mediciones repetidas como el equivalente del análisis de un diseño de bloques completos aleatorizados, donde los sujetos se consideran como los bloques.

Tabla 14-20 Análisis de varianza del diseño de mediciones repetidas con un solo factor

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F_0
1. Entre los sujetos	$\sum_{j=1}^{n} \frac{y_{.j}^2}{a} - \frac{y_{..}^2}{an}$	$n-1$		
2. Dentro de los sujetos	$\sum_{i=1}^{a}\sum_{j=1}^{n} y_{ij}^2 - \sum_{j=1}^{n} \frac{y_{.j}^2}{a}$	$n(a-1)$		
3. (Tratamientos)	$\sum_{j=1}^{a} \frac{y_{i.}^2}{n} - \frac{y_{..}^2}{an}$	$a-1$	$MS_{\text{Tratamientos}} = \frac{SS_{\text{Tratamientos}}}{a-1}$	$\frac{MS_{\text{Tratamientos}}}{MS_E}$
4. (Error)	Sustracción: línea (2) – línea (3)	$(a-1)(n-1)$	$MS_E = \frac{SS_E}{(a-1)(n-1)}$	
5. Total	$\sum_{i=1}^{a}\sum_{j=1}^{n} y_{ij}^2 - \frac{y_{..}^2}{an}$	$an-1$		

14-5 PROBLEMAS

14-1. Considere nuevamente el problema 5-22. Usar el procedimiento de Box-Cox para determinar si es apropiada (o útil) una transformación de la respuesta para analizar los datos de este experimento.

14-2. En el ejemplo 6-3 se seleccionó una transformación logarítmica para la respuesta velocidad de avance de una perforadora. Usar el procedimiento de Box-Cox para demostrar que se trata de una transformación de datos apropiada.

14-3. Considere de nuevo el experimento del proceso de fundición del problema 8-23, donde se usó un diseño factorial fraccionado 2^{6-3} para estudiar el peso del material de empaque que se adhiere a ánodos de carbono después de la cocción. Se hicieron tres réplicas de las ocho corridas del diseño, y el peso promedio y el rango de los pesos de cada combinación de prueba se trataron como las variables de respuesta. ¿Existe algún indicio de que se necesite una transformación para cualquiera de las dos respuestas?

14-4. En el problema 8-24 se usó un diseño factorial fraccionado con réplicas para estudiar el abombamiento o combadura del sustrato en la fabricación de semiconductores. Se usaron como variables de respuesta tanto la media como la desviación estándar de las mediciones de la combadura. ¿Existe algún indicio de que se necesite una transformación para cualquiera de las dos respuestas?

14-5. Considere nuevamente el experimento del recubrimiento fotoprotector del problema 8-25. Usar la varianza del espesor del recubrimiento en cada combinación de prueba como la variable de respuesta. ¿Existe algún indicio de que se necesite una transformación?

14-6. En el experimento defectos en la rejilla del problema 8-29 se empleó una variante de la transformación de la raíz cuadrada en el análisis de los datos. Usar el método de Box-Cox para determinar si ésta es la transformación apropiada.

14-7. En el diseño central compuesto del problema 11-14 se obtuvieron dos respuestas, la media y la varianza del espesor del óxido. Usar el método de Box-Cox para investigar la utilidad potencial de una transformación para estas dos respuestas. ¿Es apropiada la transformación logarítmica sugerida en el inciso *c* de ese problema?

14-8. En el diseño factorial 3^3 del problema 11-33, una de las respuestas es la desviación estándar. Usar el método de Box-Cox para investigar la utilidad de las transformaciones para esta respuesta. ¿Cambiaría su contestación si se usara la varianza como la respuesta?

14-9. En el problema 11-34 se sugiere usar $\ln(s^2)$ como la respuesta (referirse al inciso *b*). ¿El método de Box-Cox indica que es apropiada una transformación?

14-10. Un distribuidor de bebidas gaseosas está estudiando la efectividad de los métodos de descarga. Se han desarrollado tres tipos diferentes de carretillas, y se lleva a cabo un experimento en el laboratorio de ingeniería de métodos de la compañía. La variable de interés es el tiempo de descarga en minutos (y); sin embargo, el tiempo de descarga también guarda una estrecha relación con el volumen de las cajas descargadas (x). Cada carretilla se usó cuatro veces y se obtuvieron los datos siguientes. Analizar estos datos y sacar las conclusiones apropiadas. Utilizar $\alpha = 0.05$.

Tipo de carretilla					
1		2		3	
y	x	y	x	y	x
27	24	25	26	40	38
44	40	35	32	22	26
33	35	46	42	53	50
41	40	26	25	18	20

14-11. Calcular las medias ajustadas de los tratamientos y los errores estándar de éstas para los datos del problema 14-10.

14-12. A continuación se presentan las sumas de cuadrados y los productos de un análisis de covarianza de un solo factor. Terminar el análisis y sacar las conclusiones apropiadas. Utilizar $\alpha = 0.05$.

Conover, W.J. y R.L. Iman. [30b.] "Rank Transformations as a Bridge Between Parametric and Nonparametric Statistics" (con comentarios), en *The American Statistician*, vol. 35, pp. 124-133.

Conover, W.J., M.E. Johnson y M.M. Johnson. [31.] "A Comparative Study of Tests for Homogeneity of Variances, with Applications to the Outer Continental Shelf Bidding Data", en *Technometrics*, vol. 23, pp. 351-361.

Cook, D.R. [32a.] "Detection of Influential Observations in Linear Regression", en *Technometrics*, vol. 19, pp. 15-18.

Cook, D.R. [32b.] "Influential Observations in Linear Regression", en *Journal of the American Statistical Association*, vol. 74, pp. 169-174.

Cornell, J.A. [33.] *Experiments with Mixtures: Designs, Models, and the Analysis of Mixture Data*. 2a. edición, Wiley, Nueva York.

Cornfield, J. y J.W. Tukey. [34.] "Average Value of Mean Squares in Factorials", en *Annals of Mathematical Statistics*, vol. 27, pp. 907-949.

Daniel, C. [35a.] "Use of Half-Normal Plots in Interpreting Factorial Two Level Experiments", en *Technometrics*, vol. 1, pp. 311-342.

Daniel, C. [35b.] *Applications of Statistics to Industrial Experimentation*. Wiley, Nueva York.

Davies, O.L. [36.] *Design and Analysis of Industrial Experiments*. 2a. edición, Hafner Publishing Company, Nueva York.

Derringer, G. y R. Suich. [37.] "Simultaneous Optimization of Several Response Variables", en *Journal of Quality Technology*, vol. 12, pp. 214-219.

Dolby, J.L. [38.] "A Quick Method for Choosing a Transformation", en *Technometrics*, vol. 5, pp. 317-326.

Draper, N.R. y W.G. Hunter. [39.] "Transformations: Some Examples Revisited", en *Technometrics*, vol. 11, pp. 23-40.

Duncan, A.J. [40.] *Quality Control and Industrial Statistics*. 5a. edición, Richard D. Irwin, Homewood, Ill.

Duncan, D.B. [41.] "Multiple Range and Multiple F Tests", en *Biometrics*, vol. 11, pp. 1-42.

Dunnett, C.W. [42.] "New Tables for Multiple Comparisons with a Control", en *Biometrics*, vol. 20, pp. 482-491.

Eisenhart, C. [43.] "The Assumptions Underlying the Analysis of Variance", en *Biometrics*, vol. 3, pp. 1-21.

Fisher, R.A. [44a.] *Statistical Methods for Research Workers*. 13ava edición, Oliver & Boyd, Edimburgo.

Fisher, R.A. [44b.] *The Design of Experiments*. 8a. edición, Hafner Publishing Company, Nueva York.

Fisher, R.A. y F. Yates. [45.] *Statistical Tables for Biological, Agricultural, and Medical Research*. 4a. edición, Oliver & Boyd, Edimburgo.

Fries, A. y W.G. Hunter. [46.] "Minimum Aberration 2^{k-p} Designs", en *Technometrics*, vol. 22, pp. 601-608.

Gaylor, D.W. y T.D. Hartwell. [47.] "Expected Mean Squares for Nested Classifications", en *Biometrics*, vol. 25, pp. 427-430.

Gaylor, D.W. y F.N. Hopper. [48.] "Estimating the Degrees of Freedom for Linear Combinations of Mean Squares by Satterthwaite's Formula", en *Technometrics*, vol. 11, no. 4, pp. 699-706.

Graybill, F.A. y D.L. Weeks. [49.] "Combining Interblock and Intrablock Information in Balanced Incomplete Blocks", en *Annals of Mathematical Statistics*, vol. 30, pp. 799-805.

Graybill, F.A. [50.] *An Introduction to Linear Statistical Models*. Vol. 1, McGraw-Hill, Nueva York.

Graybill, F.A. y C.M. Wang. [51.] "Confidence Intervals on Nonnegative Linear Combinations of Variances", en *Journal of the American Statistical Association*, vol. 75, pp. 869-873.

Hamada, M. y N. Balakrishnan. [52.] "Analyzing Unreplicated Factorial Experiments: A Review with Some New Proposals" (con comentarios), en *Statistica Sinica*, vol. 8, pp. 1-41.

Hamada, M. y C.F.J. Wu. [53.] "Analysis of Designed Experiments with Complex Aliasing", en *Journal of Quality Technology*, vol. 24, no. 3, pp. 130-137.

Hill, W.G. y W.G. Hunter. [54.] "A Review of Response Surface Methodology: A Literature Survey", en *Technometrics*, vol. 8, pp. 571-590.

Hines, W.W. y D.C. Montgomery. [55.] *Probability and Statistics in Engineering and Management Science*. 3a. edición, Wiley, Nueva York.

Hocking, R.R. [56.] "A Discussion of the Two-Way Mixed Model", en *The American Statistician*, vol. 27, no. 4, pp. 148-152.

Hocking, R.R., O.P. Hackney y F.M. Speed. [57.] "The Analysis of Linear Models with Unbalanced Data", en *Contributions to Survey Sampling and Applied Statistics*, H.A. David (ed.), Academic Press, Nueva York.

Hocking, R.R. y F.M. Speed. [58.] "A Full Rank Analysis of Some Linear Model Problems", en *Journal of the American Statistical Association*, vol. 70, pp. 706-712.

Hunter, J.S. [59a.] "Statistical Design Applied to Product Design", en *Journal of Quality Technology*, vol. 17, pp. 210-221.

Hunter, J.S. [59b.] "Let's All Beware the Latin Square", en *Quality Engineering*, vol. 1, pp. 453-465.

John, J.A. y P. Prescott. [60.] "Critical Values of a Test to Detect Outliers in Factorial Experiments", en *Applied Statistics*, vol. 24, pp. 56-59.

John, P.W.M. [61a.] "The Three-Quarter Replicates of 2^4 and 2^5 Designs", en *Biometrics*, vol. 17, pp. 319-321.

John, P.W.M. [61b.] "Three-Quarter Replicates of 2^n Designs", en *Biometrics*, vol. 18, pp. 171-184.

John, P.W.M. [61c.] "Blocking a $3(2^{n-k})$ Design", en *Technometrics*, vol. 6, pp. 371-376.

John, P.W.M. [61d.] *Statistical Design and Analysis of Experiments*. Macmillan, Nueva York.

Kackar, R.N. [62.] "Off-Line Quality Control, Parameter Design, and the Taguchi Method", en *Journal of Quality Technology*, vol. 17, pp. 176-188.

Kempthorne, O. [63.] *The Design and Analysis of Experiments*. Wiley, Nueva York.

Keuls, M. [64.] "The Use of the Studentized Range in Connection with an Analysis of Variance", en *Euphytica*, vol. 1, pp. 112-122.

Kiefer, J. [65a.] "Optimum Experimental Designs", en *Journal of the Royal Statistical Society B*, vol. 21, pp. 272-304.

Kiefer, J. [65b.] "Optimum Designs in Regression Problems", en *Annals of Mathematical Statistics*, vol. 32, pp. 298-325.

Kiefer, J. y J. Wolfowitz. [66.] "Optimum Designs in Regression Problems", en *Annals of Mathematical Statistics*, vol. 30, pp. 271-294.

Khuri, A.I. y J.A. Cornell. [67.] *Response Surfaces: Designs and Analyses*. 2a. edición, Dekker, Nueva York.

Kruskal, W.H. y W.A. Wallis. [68.] "Use of Ranks on One Criterion Variance Analysis", en *Journal of the American Statistical Association*, vol. 47, pp. 583-621 (las correcciones aparecen en el vol. 48, pp. 907-911).

Larntz, K. y P. Whitcomb. [69.] "Use of Replication in Almost Unreplicated Factorials", ponencia presentada en la Conferencia Técnica de Otoño, Corning, N.Y.

Lenth, R.V. [70.] "Quick and Easy Analysis of Unreplicated Factorials", en *Technometrics*, vol. 31, pp. 469-473.

Leon, R.V., A.C. Shoemaker y R.N. Kackar. [71.] "Performance Measures Independent of Adjustment", en *Technometrics*, vol. 29, pp. 253-265.

Levene, H. [72.] "Robust Tests for Equality of Variance", en *Contributions to Probability and Statistics*, Z. Olkin, ed., Stanford University Press, Palo Alto, CA, pp. 278-292.

Loughin, T.M. [73.] "Calibration of the Lenth Test for Unreplicated Factorial Designs", en *Journal of Quality Technology*, vol. 30, pp. 171-175.

Loughin, T.M. y W. Noble. [74.] "A Permutation Test for Effects in an Unreplicated Factorial Design", en *Technometrics*, vol. 39, pp. 180-190.

Margolin, B.H. [75a.] "Systematic Methods of Analyzing 2^n3^m Factorial Experiments with Applications", en *Technometrics*, vol. 9, pp. 245-260.

Margolin, B.H. [75b.] "Results on Factorial Designs of Resolution IV for the 2^n and 2^n3^m Series", en *Technometrics*, vol. 11, pp. 431-444.

McCullagh, P. y J.A. Nelder. [76.] *Generalized Linear Models*. 2a. edición, Chapman & Hall, Nueva York.

Miller, R.G. [77.] *Simultaneous Statistical Inference*. Springer-Verlag, Nueva York.

Miller, R.G., Jr. [78.] "Developments in Multiple Comparisons, 1966-1976", en *Journal of the American Statistical Association*, vol. 72, pp. 779-788.

Milliken, G.A. y D.E. Johnson. [79.] *Analysis of Messy Data, Vol. 1*. Van Nostrand Reinhold, Nueva York.

Montgomery, D.C. [80a.] *Introduction to Statistical Quality Control*. 3a. edición, Wiley, Nueva York.

Montgomery, D.C. [80b.] "Experimental Design for Product and Process Design and Development", en *Journal of the Royal Statistical Society*, D, vol. 48, pp. 159-177.

Montgomery, D.C., C.M. Borror y J.D. Stanley. [81.] "Some Cautions in the Use of Plackett-Burman Designs", en *Quality Engineering*, vol. 10, pp. 371-381.

Montgomery, D.C. y E.A. Peck. [82.] *Introduction to Linear Regression Analysis*. 2a. edición, Wiley, Nueva York.

Montgomery, D.C. y G.C. Runger. [83a.] "Gauge Capability Analysis and Designed Experiments. Part I: Basic Methods", en *Quality Engineering*, vol. 6, pp. 115-135.

Montgomery, D.C. y G.C. Runger. [83b.] "Gauge Capability Analysis and Designed Experiments. Part II: Experimental Design Models and Variance Component Estimation", en *Quality Engineering*, vol. 6, pp. 289-305.

Montgomery, D.C. y G.C. Runger. [83c.] "Foldovers of 2^{k-p} Resolution IV Designs", en *Journal of Quality Technology*, vol. 24, pp. 446-450.

Montgomery, D.C. y G.C. Runger. [83d.] *Probabilidad y estadística aplicadas para ingenieros*. 2a. edición, Limusa-Wiley, México.

Myers, R.H. [84.] *Classical and Modern Regression with Applications*. 2a. edición, PNS-Kent, Boston.

Myers, R.H. y D.C. Montgomery. [85a.] *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*, Wiley, Nueva York.

Myers, R.H. y D.C. Montgomery. [85b.] "A Tutorial on Generalized Linear Models", en *Journal of Quality Technology*, vol. 29, pp. 274-291.

Nair, V.N. *et al.* (eds.) [86.] "Taguchi's Parameter Design: A Panel Discussion", en *Technometrics*, vol. 34, pp. 127-161.

Nelder, J.A. y R.W.M. Wedderburn. [87.] "Generalized Linear Models", en *Journal of the Royal Statistical Society*, A, vol. 135, pp. 370-384.

Nelson, L.S. [88a.] "Using Nested Designs I: Estimation of Standard Deviations", en *Journal of Quality Technology*, vol. 27, no. 2, pp. 169-171.

Nelson, L.S. [88b.] "Using Nested Designs II: Confidence Limits for Standard Deviations", en *Journal of Quality Technology*, vol. 27, no. 3, pp. 265-267.

Nelson, L.S. [88c.] "Variance Estimation Using Staggered, Nested Designs", en *Journal of Quality Technology*, vol. 15, pp. 195-198.

Nelson, P.R. [89.] "Multiple Comparison of Means Using Simultaneous Confidence Intervals", en *Journal of Quality Technology*, vol. 21, No. 4, pp. 232-241.

Newman, D. [90.] "The Distribution of the Range in Samples from a Normal Population, Expressed in Terms of an Independent Estimate of Standard Deviation", en *Biometrika*, vol. 31, pp. 20-30.

O'Neill, R. y G.B. Wetherill. [91.] "The Present State of Multiple Comparison Methods", en *Journal of the Royal Statistical Society*, B, vol. 33, pp. 218-241.

Ostle, B. [92.] *Statistics in Research*. 2a. edición, Iowa State Press, Ames, Iowa.

Pearson, E.S. y H.O. Hartley. [93a.] *Biometrika Tables for Statisticians*. Vol. 1, 3a. edición, Cambridge University Press, Cambridge.

Pearson, E.S. y H.O. Hartley. [93b.] *Biometrika Tables for Statisticians*. Vol. 2, Cambridge University Press, Cambridge.

Pignatiello, J.J., Jr. y J.S. Ramberg. [94.] "Top Ten Triumphs and Tragedies of Genichi Taguchi", en *Quality Engineering*, vol. 4, pp. 211-225.

Plackett, R.L. y J.P. Burman. [95.] "The Design of Optimum Multifactorial Experiments", en *Biometrika*, vol. 33, pp. 305-325.

Quenouille, M.H. [96.] *The Design and Analysis of Experiments*. Charles Griffin & Company, Londres.

Satterthwaite, F.E. [97.] "An Approximate Distribution of Estimates of Variance Components", en *Biometrics Bull.*, vol. 2, pp. 110-112.

Scheffé, H. [98a.] "A Method for Judging All Contrasts in the Analysis of Variance", en *Biometrika*, vol. 40, pp. 87-104.

Scheffé, H. [98b.] "A 'Mixed Model' for the Analysis of Variance", en *Annals of Mathematical Statistics*, vol. 27, pp. 23-36.

Scheffé, H. [98c.] "Alternative Models for the Analysis of Variance", en *Annals of Mathematical Statistics*, vol. 27, pp. 251-271.

Scheffé, H. [98d.] *The Analysis of Variance*. Wiley, Nueva York.

Searle, S.R. [99a.] *Linear Models*. Wiley, Nueva York.

Searle, S.R. [99b.] "Topics in Variance Component Estimation", en *Biometrics*, vol. 27, pp. 1-76.

Searle, S.R. [99c.] *Linear Models for Unbalanced Data*. Wiley, Nueva York.

Searle, S.R., G. Casella y G.E. McCulloch. [100.] *Variance Components*. Wiley, Nueva York.

Searle, S.R. y R.F. Fawcett. [101.] "Expected Mean Squares in Variance Component Models Having Finite Populations", en *Biometrics*, vol. 26, pp. 243-254.

Searle, S.R., F.M. Speed y H.V. Henderson. [102.] "Some Computational and Model Equivalences in Analyses of Variance of Unequal-Subclass-Numbers Data", en *The American Statistician*, vol. 35, pp. 16-33.

Smith, H.F. [103.] "Interpretations of Adjusted Treatment Means and Regressions in Analysis of Covariance", en *Biometrics*, vol. 13, No. 3, pp. 282-308.

Smith, J.R. y J.M. Beverly. [104.] "The Use and Analysis of Staggered Nested Factorial Designs", en *Journal of Quality Technology*, vol. 13, pp. 166-173.

Speed, F.M. y R.R. Hocking. [105.] "The Use of the *R*( )-Notation with Unbalanced Data", en *The American Statistician*, vol. 30, pp. 30-33.

Speed, F.M., R.R. Hocking y O.P. Hackney. [106.] "Methods of Analysis of Linear Models with Unbalanced Data", en *Journal of the American Statistical Association*, vol. 73, pp. 105-112.

Stefansky, W. [107.] "Rejecting Outliers in Factorial Designs", en *Technometrics*, vol. 14, pp. 469-479.

Taguchi, G. [108a.] *System of Experimental Design: Engineering Methods to Optimize Quality and Minimize Cost*. UNIPUB, White Plains, Nueva York.

Taguchi, G. [108b.] *Introduction to Quality Engineering*. Asian Productivity Organization, UNIPUB, White Plains, Nueva York.

Taguchi, G. y Y. Wu. [109.] *Introduction to Off-Line Quality Control*. Central Japan Quality Control Association, Nagoya, Japón.

Ting, N., R.K. Burdick, F.A. Graybill, S. Jeyaratnam y T.-F. C. Lu. [110.] "Confidence Intervals on Linear Combinations of Variance Components That Are Unrestricted in Sign", en *Journal of Statistical Computation and Simulation*, vol. 35, pp. 135-143.

Tukey, J.W. [111a.] "One Degree of Freedom for Non-Additivity", en *Biometrics*, vol. 5, pp. 232-242.

Tukey, J.W. [111b.] "Comparing Individual Means in the Analysis of Variance", en *Biometrics*, vol. 5, pp. 99-114.

Tukey, J.W. [111c.] "Quick and Dirty Methods in Statistics, Part II, Simple Analysis for Standard Designs", en *Proceedings of the Fifth Annual Convention, American Society for Quality Control*, pp. 189-197.

Tukey, J.W. [111d.] "The Problem of Multiple Comparisons", notas inéditas, Princeton University.

Winer, B.J. [112.] *Statistical Principles in Experimental Design*. 2a. edición, McGraw-Hill, Nueva York.

Yates, F. [113a.] "The Analysis of Multiple Classifications with Unequal Numbers in the Different Classes", en *Journal of the American Statistical Association*, vol. 29, pp. 52-66.

Yates, F. [113b.] *Design and Analysis of Factorial Experiments*. Comunicado técnico No. 35, Imperial Bureau of Soil Sciences, Londres.

Yates, F. [113c.] "The Recovery of Interblock Information in Balanced Incomplete Block Designs", en *Annals of Eugenics*, vol. 10, pp. 317-325.

Ye, K. y M. Hamada. [114.] "Critical Values of the Lenth Method for Unreplicated Factorial Designs", en *Journal of Quality Technology*, vol. 32, pp. 57-66.

Apéndice

- **I** Distribución normal estándar acumulada
- **II** Puntos porcentuales de la distribución t
- **III** Puntos porcentuales de la distribución χ^2
- **IV** Puntos porcentuales de la distribución F
- **V** Curvas de operación característica para el análisis de varianza del modelo con efectos fijos
- **VI** Curvas de operación característica para el análisis de varianza del modelo con efectos aleatorios
- **VII** Rangos significativos para la prueba del rango múltiple de Duncan
- **VIII** Puntos porcentuales del estadístico del rango studentizado
- **IX** Valores críticos para la prueba de Dunnett para comparar tratamientos con un control
- **X** Coeficientes de polinomios ortogonales
- **XI** Números aleatorios
- **XII** Relaciones de alias para diseños factoriales fraccionados 2^{k-p} con $k \leq 15$ y $n \leq 64$
- **XIII** Glosario para el uso de *Design Expert*

I. Distribución normal estándar acumulada[a]

$$\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-u^2/2}\, du$$

z	.00	.01	.02	.03	.04	z
.0	.50000	.50399	.50798	.51197	.51595	0
.1	.53983	.54379	.54776	.55172	.55567	.1
.2	.57926	.58317	.58706	.59095	.59483	.2
.3	.61791	.62172	.62551	.62930	.63307	.3
.4	.65542	.65910	.66276	.66640	.67003	.4
.5	.69146	.69497	.69847	.70194	.70540	.5
.6	.72575	.72907	.73237	.73565	.73891	.6
.7	.75803	.76115	.76424	.76730	.77035	.7
.8	.78814	.79103	.79389	.79673	.79954	.8
.9	.81594	.81859	.82121	.82381	.82639	.9
1.0	.84134	.84375	.84613	.84849	.85083	1.0
1.1	.86433	.86650	.86864	.87076	.87285	1.1
1.2	.88493	.88686	.88877	.89065	.89251	1.2
1.3	.90320	.90490	.90658	.90824	.90988	1.3
1.4	.91924	.92073	.92219	.92364	.92506	1.4
1.5	.93319	.93448	.93574	.93699	.93822	1.5
1.6	.94520	.94630	.94738	.94845	.94950	1.6
1.7	.95543	.95637	.95728	.95818	.95907	1.7
1.8	.96407	.96485	.96562	.96637	.96711	1.8
1.9	.97128	.97193	.97257	.97320	.97381	1.9
2.0	.97725	.97778	.97831	.97882	.97932	2.0
2.1	.98214	.98257	.98300	.98341	.93882	2.1
2.2	.98610	.98645	.98679	.98713	.98745	2.2
2.3	.98928	.98956	.98983	.99010	.99036	2.3
2.4	.99180	.99202	.99224	.99245	.99266	2.4
2.5	.99379	.99396	.99413	.99430	.99446	2.5
2.6	.99534	.99547	.99560	.99573	.99585	2.6
2.7	.99653	.99664	.99674	.99683	.99693	2.7
2.8	.99744	.99752	.99760	.99767	.99774	2.8
2.9	.99813	.99819	.99825	.99831	.99836	2.9
3.0	.99865	.99869	.99874	.99878	.99882	3.0
3.1	.99903	.99906	.99910	.99913	.99916	3.1
3.2	.99931	.99934	.99936	.99938	.99940	3.2
3.3	.99952	.99953	.99955	.99957	.99958	3.3
3.4	.99966	.99968	.99969	.99970	.99971	3.4
3.5	.99977	.99978	.99978	.99979	.99980	3.5
3.6	.99984	.99985	.99985	.99986	.99986	3.6
3.7	.99989	.99990	.99990	.99990	.99991	3.7
3.8	.99993	.99993	.99993	.99994	.99994	3.8
3.9	.99995	.99995	.99996	.99996	.99996	3.9

[a]Reproducida con permiso de *Probability and Statistics in Engineering and Management Science*, 3a. ed., W.W. Hines y D.C. Montgomery, Wiley, Nueva York

I. Distribución normal estándar acumulada (*continuación*)

$$\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-u^2/2}\, du$$

z	.05	.06	.07	.08	.09	*z*
.0	.51994	.52392	.52790	.53188	.53586	.0
.1	.55962	.56356	.56749	.57142	.57534	.1
.2	.59871	.60257	.60642	.61026	.61409	.2
.3	.63683	.64058	.64431	.64803	.65173	.3
.4	.67364	.67724	.68082	.68438	.68793	.4
.5	.70884	.71226	.71566	.71904	.72240	.5
.6	.74215	.74537	.74857	.75175	.75490	.6
.7	.77337	.77637	.77935	.78230	.78523	.7
.8	.80234	.80510	.80785	.81057	.81327	.8
.9	.82894	.83147	.83397	.83646	.83891	.9
1.0	.85314	.85543	.85769	.85993	.86214	1.0
1.1	.87493	.87697	.87900	.88100	.88297	1.1
1.2	.89435	.89616	.89796	.89973	.90147	1.2
1.3	.91149	.91308	.91465	.91621	.91773	1.3
1.4	.92647	.92785	.92922	.93056	.93189	1.4
1.5	.93943	.90462	.94179	.94295	.94408	1.5
1.6	.95053	.95154	.95254	.95352	.95448	1.6
1.7	.95994	.96080	.96164	.96246	.96327	1.7
1.8	.96784	.96856	.96926	.96995	.97062	1.8
1.9	.97441	.97500	.97558	.97615	.97670	1.9
2.0	.97982	.98030	.98077	.98124	.98169	2.0
2.1	.98422	.98461	.98500	.98537	.98574	2.1
2.2	.98778	.98809	.98840	.98870	.98899	2.2
2.3	.99061	.99086	.99111	.99134	.99158	2.3
2.4	.99286	.99305	.99324	.99343	.99361	2.4
2.5	.99461	.99477	.99492	.99506	.99520	2.5
2.6	.99598	.99609	.99621	.99632	.99643	2.6
2.7	.99702	.99711	.99720	.99728	.99736	2.7
2.8	.99781	.99788	.99795	.99801	.99807	2.8
2.9	.99841	.99846	.99851	.99856	.99861	2.9
3.0	.99886	.99889	.99893	.99897	.99900	3.0
3.1	.99918	.99921	.99924	.99926	.99929	3.1
3.2	.99942	.99944	.99946	.99948	.99950	3.2
3.3	.99960	.99961	.99962	.99964	.99965	3.3
3.4	.99972	.99973	.99974	.99975	.99976	3.4
3.5	.99981	.99981	.99982	.99983	.99983	3.5
3.6	.99987	.99987	.99988	.99988	.99989	3.6
3.7	.99991	.99992	.99992	.99992	.99992	3.7
3.8	.99994	.99994	.99995	.99995	.99995	3.8
3.9	.99996	.99996	.99996	.99997	.99997	3.9

II. Puntos porcentuales de la distribución t[a]

ν \ α	.40	.25	.10	.05	.025	.01	.005	.0025	.001	.0005
1	.325	1.000	3.078	6.314	12.706	31.821	63.657	127.32	318.31	636.62
2	.289	.816	1.886	2.920	4.303	6.965	9.925	14.089	23.326	31.598
3	.277	.765	1.638	2.353	3.182	4.541	5.841	7.453	10.213	12.924
4	.271	.741	1.533	2.132	2.776	3.747	4.604	5.598	7.173	8.610
5	.267	.727	1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.869
6	.265	.727	1.440	1.943	2.447	3.143	3.707	4.317	5.208	5.959
7	.263	.711	1.415	1.895	2.365	2.998	3.499	4.019	4.785	5.408
8	.262	.706	1.397	1.860	2.306	2.896	3.355	3.833	4.501	5.041
9	.261	.703	1.383	1.833	2.262	2.821	3.250	3.690	4.297	4.781
10	.260	.700	1.372	1.812	2.228	2.764	3.169	3.581	4.144	4.587
11	.260	.697	1.363	1.796	2.201	2.718	3.106	3.497	4.025	4.437
12	.259	.695	1.356	1.782	2.179	2.681	3.055	3.428	3.930	4.318
13	.259	.694	1.350	1.771	2.160	2.650	3.012	3.372	3.852	4.221
14	.258	.692	1.345	1.761	2.145	2.624	2.977	3.326	3.787	4.140
15	.258	.691	1.341	1.753	2.131	2.602	2.947	3.286	3.733	4.073
16	.258	.690	1.337	1.746	2.120	2.583	2.921	3.252	3.686	4.015
17	.257	.689	1.333	1.740	2.110	2.567	2.898	3.222	3.646	3.965
18	.257	.688	1.330	1.734	2.101	2.552	2.878	3.197	3.610	3.922
19	.257	.688	1.328	1.729	2.093	2.539	2.861	3.174	3.579	3.883
20	.257	.687	1.325	1.725	2.086	2.528	2.845	3.153	3.552	3.850
21	.257	.686	1.323	1.721	2.080	2.518	2.831	3.135	3.527	3.819
22	.256	.686	1.321	1.717	2.074	2.508	2.819	3.119	3.505	3.792
23	.256	.685	1.319	1.714	2.069	2.500	2.807	3.104	3.485	3.767
24	.256	.685	1.318	1.711	2.064	2.492	2.797	3.091	3.467	3.745
25	.256	.684	1.316	1.708	2.060	2.485	2.787	3.078	3.450	3.725
26	.256	.684	1.315	1.706	2.056	2.479	2.779	3.067	3.435	3.707
27	.256	.684	1.314	1.703	2.052	2.473	2.771	3.057	3.421	3.690
28	.256	.683	1.313	1.701	2.048	2.467	2.763	3.047	3.408	3.674
29	.256	.683	1.311	1.699	2.045	2.462	2.756	3.038	3.396	3.659
30	.256	.683	1.310	1.697	2.042	2.457	2.750	3.030	3.385	3.646
40	.255	.681	1.303	1.684	2.021	2.423	2.704	2.971	3.307	3.551
60	.254	.679	1.296	1.671	2.000	2.390	2.660	2.915	3.232	3.460
120	.254	.677	1.289	1.658	1.980	2.358	2.617	2.860	3.160	3.373
∞	.253	.674	1.282	1.645	1.960	2.326	2.576	2.807	3.090	3.291

ν = grados de libertad.

[a]Adaptada con permiso de *Biometrika Tables for Statisticians*, vol. 1, 3a. ed., E.S. Pearson y H.O. Hartley, Cambridge University Press, Cambridge

III. Puntos porcentuales de la distribución χ^{2} [a]

ν	.995	.990	.975	.950	.500	.050	.025	.010	.005
1	0.00 +	0.00 +	0.00 +	0.00 +	0.45	3.84	5.02	6.63	7.88
2	0.01	0.02	0.05	0.10	1.39	5.99	7.38	9.21	10.60
3	0.07	0.11	0.22	0.35	2.37	7.81	9.35	11.34	12.84
4	0.21	0.30	0.48	0.71	3.36	9.49	11.14	13.28	14.86
5	0.41	0.55	0.83	1.15	4.35	11.07	12.38	15.09	16.75
6	0.68	0.87	1.24	1.64	5.35	12.59	14.45	16.81	18.55
7	0.99	1.24	1.69	2.17	6.35	14.07	16.01	18.48	20.28
8	1.34	1.65	2.18	2.73	7.34	15.51	17.53	20.09	21.96
9	1.73	2.09	2.70	3.33	8.34	16.92	19.02	21.67	23.59
10	2.16	2.56	3.25	3.94	9.34	18.31	20.48	23.21	25.19
11	2.60	3.05	3.82	4.57	10.34	19.68	21.92	24.72	26.76
12	3.07	3.57	4.40	5.23	11.34	21.03	23.34	26.22	28.30
13	3.57	4.11	5.01	5.89	12.34	22.36	24.74	27.69	29.82
14	4.07	4.66	5.63	6.57	13.34	23.68	26.12	29.14	31.32
15	4.60	5.23	6.27	7.26	14.34	25.00	27.49	30.58	32.80
16	5.14	5.81	6.91	7.96	15.34	26.30	28.85	32.00	34.27
17	5.70	6.41	7.56	8.67	16.34	27.59	30.19	33.41	35.72
18	6.26	7.01	8.23	9.39	17.34	28.87	31.53	34.81	37.16
19	6.84	7.63	8.91	10.12	18.34	30.14	32.85	36.19	38.58
20	7.43	8.26	9.59	10.85	19.34	31.41	34.17	37.57	40.00
25	10.52	11.52	13.12	14.61	24.34	37.65	40.65	44.31	46.93
30	13.79	14.95	16.79	18.49	29.34	43.77	46.98	50.89	53.67
40	20.71	22.16	24.43	26.51	39.34	55.76	59.34	63.69	66.77
50	27.99	29.71	32.36	34.76	49.33	67.50	71.42	76.15	79.49
60	35.53	37.48	40.48	43.19	59.33	79.08	83.30	88.38	91.95
70	43.28	45.44	48.76	51.74	69.33	90.53	95.02	100.42	104.22
80	51.17	53.54	57.15	60.39	79.33	101.88	106.63	112.33	116.32
90	59.20	61.75	65.65	69.13	89.33	113.14	118.14	124.12	128.30
100	67.33	70.06	74.22	77.93	99.33	124.34	129.56	135.81	140.17

(The column headings are values of α.)

ν = grados de libertad

[a]Adaptada con permiso de *Biometrika Tables for Statisticians*, vol. 1, 3a. ed., E.S. Pearson y H.O. Hartley, Cambridge University Press, Cambridge

IV. Puntos porcentuales de la distribución F[a]

$F_{0.25, \nu_1, \nu_2}$

ν_2 \ ν_1	Grados de libertad del numerador (ν_1) 1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	5.83	7.50	8.20	8.58	8.82	8.98	9.10	9.19	9.26	9.32	9.41	9.49	9.58	9.63	9.67	9.71	9.76	9.80	9.85
2	2.57	3.00	3.15	3.23	3.28	3.31	3.34	3.35	3.37	3.38	3.39	3.41	3.43	3.43	3.44	3.45	3.46	3.47	3.48
3	2.02	2.28	2.36	2.39	2.41	2.42	2.43	2.44	2.44	2.44	2.45	2.46	2.46	2.46	2.47	2.47	2.47	2.47	2.47
4	1.81	2.00	2.05	2.06	2.07	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08
5	1.69	1.85	1.88	1.89	1.89	1.89	1.89	1.89	1.89	1.89	1.89	1.89	1.88	1.88	1.88	1.88	1.87	1.87	1.87
6	1.62	1.76	1.78	1.79	1.79	1.78	1.78	1.78	1.77	1.77	1.77	1.76	1.76	1.75	1.75	1.75	1.74	1.74	1.74
7	1.57	1.70	1.72	1.72	1.71	1.71	1.70	1.70	1.70	1.69	1.68	1.68	1.67	1.67	1.66	1.66	1.65	1.65	1.65
8	1.54	1.66	1.67	1.66	1.66	1.65	1.64	1.64	1.63	1.63	1.62	1.62	1.61	1.60	1.60	1.59	1.59	1.58	1.58
9	1.51	1.62	1.63	1.63	1.62	1.61	1.60	1.60	1.59	1.59	1.58	1.57	1.56	1.56	1.55	1.54	1.54	1.53	1.53
10	1.49	1.60	1.60	1.59	1.59	1.58	1.57	1.56	1.56	1.55	1.54	1.53	1.52	1.52	1.51	1.51	1.50	1.49	1.48
11	1.47	1.58	1.58	1.57	1.56	1.55	1.54	1.53	1.53	1.52	1.51	1.50	1.49	1.49	1.48	1.47	1.47	1.46	1.45
12	1.46	1.56	1.56	1.55	1.54	1.53	1.52	1.51	1.51	1.50	1.49	1.48	1.47	1.46	1.45	1.45	1.44	1.43	1.42
13	1.45	1.55	1.55	1.53	1.52	1.51	1.50	1.49	1.49	1.48	1.47	1.46	1.45	1.44	1.43	1.42	1.42	1.41	1.40
14	1.44	1.53	1.53	1.52	1.51	1.50	1.49	1.48	1.47	1.46	1.45	1.44	1.43	1.42	1.41	1.41	1.40	1.39	1.38
15	1.43	1.52	1.52	1.51	1.49	1.48	1.47	1.46	1.46	1.45	1.44	1.43	1.41	1.41	1.40	1.39	1.38	1.37	1.36
16	1.42	1.51	1.51	1.50	1.48	1.47	1.46	1.45	1.44	1.44	1.43	1.41	1.40	1.39	1.38	1.37	1.36	1.35	1.34
17	1.42	1.51	1.50	1.49	1.47	1.46	1.45	1.44	1.43	1.43	1.41	1.40	1.39	1.38	1.37	1.36	1.35	1.34	1.33
18	1.41	1.50	1.49	1.48	1.46	1.45	1.44	1.43	1.42	1.42	1.40	1.39	1.38	1.37	1.36	1.35	1.34	1.33	1.32
19	1.41	1.49	1.49	1.47	1.46	1.44	1.43	1.42	1.41	1.41	1.40	1.38	1.37	1.36	1.35	1.34	1.33	1.32	1.30
20	1.40	1.49	1.48	1.47	1.45	1.44	1.43	1.42	1.41	1.40	1.39	1.37	1.36	1.35	1.34	1.33	1.32	1.31	1.29
21	1.40	1.48	1.48	1.46	1.44	1.43	1.42	1.41	1.40	1.39	1.38	1.37	1.35	1.34	1.33	1.32	1.31	1.30	1.28
22	1.40	1.48	1.47	1.45	1.44	1.42	1.41	1.40	1.39	1.39	1.37	1.36	1.34	1.33	1.32	1.31	1.30	1.29	1.28
23	1.39	1.47	1.47	1.45	1.43	1.42	1.41	1.40	1.39	1.38	1.37	1.35	1.34	1.33	1.32	1.31	1.30	1.28	1.27
24	1.39	1.47	1.46	1.44	1.43	1.41	1.40	1.39	1.38	1.38	1.36	1.35	1.33	1.32	1.31	1.30	1.29	1.28	1.26
25	1.39	1.47	1.46	1.44	1.42	1.41	1.40	1.39	1.38	1.37	1.36	1.34	1.33	1.32	1.31	1.29	1.28	1.27	1.25
26	1.38	1.46	1.45	1.44	1.42	1.41	1.39	1.38	1.37	1.37	1.35	1.34	1.32	1.31	1.30	1.29	1.28	1.26	1.25
27	1.38	1.46	1.45	1.43	1.42	1.40	1.39	1.38	1.37	1.36	1.35	1.33	1.32	1.31	1.30	1.28	1.27	1.26	1.24
28	1.38	1.46	1.45	1.43	1.41	1.40	1.39	1.38	1.37	1.36	1.34	1.33	1.31	1.30	1.29	1.28	1.27	1.25	1.24
29	1.38	1.45	1.45	1.43	1.41	1.40	1.38	1.37	1.36	1.35	1.34	1.32	1.31	1.30	1.29	1.27	1.26	1.25	1.23
30	1.38	1.45	1.44	1.42	1.41	1.39	1.38	1.37	1.36	1.35	1.34	1.32	1.30	1.29	1.28	1.27	1.26	1.24	1.23
40	1.36	1.44	1.42	1.40	1.39	1.37	1.36	1.35	1.34	1.33	1.31	1.30	1.28	1.26	1.25	1.24	1.22	1.21	1.19
60	1.35	1.42	1.41	1.38	1.37	1.35	1.33	1.32	1.31	1.30	1.29	1.27	1.25	1.24	1.22	1.21	1.19	1.17	1.15
120	1.34	1.40	1.39	1.37	1.35	1.33	1.31	1.30	1.29	1.28	1.26	1.24	1.22	1.21	1.19	1.18	1.16	1.13	1.10
∞	1.32	1.39	1.37	1.35	1.33	1.31	1.29	1.28	1.27	1.25	1.24	1.22	1.19	1.18	1.16	1.14	1.12	1.08	1.00

Grados de libertad del denominador (ν_2)

ν = grados de libertad

[a]Adaptada con permiso de *Biometrika Tables for Statisticians*, vol. 1, 3a. ed., E.S. Pearson y H.O. Hartley, Cambridge University Press, Cambridge

$F_{0.10, \nu_1, \nu_2}$

Grados de libertad del denominador (ν_2) \ Grados de libertad del numerador (ν_1)	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	39.86	49.50	53.59	55.83	57.24	58.20	58.91	59.44	59.86	60.19	60.71	61.22	61.74	62.00	62.26	62.53	62.79	63.06	63.33
2	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38	9.39	9.41	9.42	9.44	9.45	9.46	9.47	9.47	9.48	9.49
3	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24	5.23	5.22	5.20	5.18	5.18	5.17	5.16	5.15	5.14	5.13
4	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94	3.92	3.90	3.87	3.84	3.83	3.82	3.80	3.79	3.78	3.76
5	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32	3.30	3.27	3.24	3.21	3.19	3.17	3.16	3.14	3.12	3.10
6	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96	2.94	2.90	2.87	2.84	2.82	2.80	2.78	2.76	2.74	2.72
7	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72	2.70	2.67	2.63	2.59	2.58	2.56	2.54	2.51	2.49	2.47
8	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56	2.54	2.50	2.46	2.42	2.40	2.38	2.36	2.34	2.32	2.29
9	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44	2.42	2.38	2.34	2.30	2.28	2.25	2.23	2.21	2.18	2.16
10	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35	2.32	2.28	2.24	2.20	2.18	2.16	2.13	2.11	2.08	2.06
11	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.30	2.27	2.25	2.21	2.17	2.12	2.10	2.08	2.05	2.03	2.00	1.97
12	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21	2.19	2.15	2.10	2.06	2.04	2.01	1.99	1.96	1.93	1.90
13	3.14	2.76	2.56	2.43	2.35	2.28	2.23	2.20	2.16	2.14	2.10	2.05	2.01	1.98	1.96	1.93	1.90	1.88	1.85
14	3.10	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.12	2.10	2.05	2.01	1.96	1.94	1.91	1.89	1.86	1.83	1.80
15	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09	2.06	2.02	1.97	1.92	1.90	1.87	1.85	1.82	1.79	1.76
16	3.05	2.67	2.46	2.33	2.24	2.18	2.13	2.09	2.06	2.03	1.99	1.94	1.89	1.87	1.84	1.81	1.78	1.75	1.72
17	3.03	2.64	2.44	2.31	2.22	2.15	2.10	2.06	2.03	2.00	1.96	1.91	1.86	1.84	1.81	1.78	1.75	1.72	1.69
18	3.01	2.62	2.42	2.29	2.20	2.13	2.08	2.04	2.00	1.98	1.93	1.89	1.84	1.81	1.78	1.75	1.72	1.69	1.66
19	2.99	2.61	2.40	2.27	2.18	2.11	2.06	2.02	1.98	1.96	1.91	1.86	1.81	1.79	1.76	1.73	1.70	1.67	1.63
20	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96	1.94	1.89	1.84	1.79	1.77	1.74	1.71	1.68	1.64	1.61
21	2.96	2.57	2.36	2.23	2.14	2.08	2.02	1.98	1.95	1.92	1.87	1.83	1.78	1.75	1.72	1.69	1.66	1.62	1.59
22	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93	1.90	1.86	1.81	1.76	1.73	1.70	1.67	1.64	1.60	1.57
23	2.94	2.55	2.34	2.21	2.11	2.05	1.99	1.96	1.92	1.89	1.84	1.80	1.74	1.72	1.69	1.66	1.62	1.59	1.55
24	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91	1.88	1.83	1.78	1.73	1.70	1.67	1.64	1.61	1.57	1.53
25	2.92	2.53	2.32	2.18	2.09	2.02	1.97	1.93	1.89	1.87	1.82	1.77	1.72	1.69	1.66	1.63	1.59	1.56	1.52
26	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.88	1.86	1.81	1.76	1.71	1.68	1.65	1.61	1.58	1.54	1.50
27	2.90	2.51	2.30	2.17	2.07	2.00	1.95	1.91	1.87	1.85	1.80	1.75	1.70	1.67	1.64	1.60	1.57	1.53	1.49
28	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87	1.84	1.79	1.74	1.69	1.66	1.63	1.59	1.56	1.52	1.48
29	2.89	2.50	2.28	2.15	2.06	1.99	1.93	1.89	1.86	1.83	1.78	1.73	1.68	1.65	1.62	1.58	1.55	1.51	1.47
30	2.88	2.49	2.28	2.14	2.03	1.98	1.93	1.88	1.85	1.82	1.77	1.72	1.67	1.64	1.61	1.57	1.54	1.50	1.46
40	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79	1.76	1.71	1.66	1.61	1.57	1.54	1.51	1.47	1.42	1.38
60	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74	1.71	1.66	1.60	1.54	1.51	1.48	1.44	1.40	1.35	1.29
120	2.75	2.35	2.13	1.99	1.90	1.82	1.77	1.72	1.68	1.65	1.60	1.55	1.48	1.45	1.41	1.37	1.32	1.26	1.19
∞	2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.63	1.60	1.55	1.49	1.42	1.38	1.34	1.30	1.24	1.17	1.00

IV. Puntos porcentuales de la distribución F (*continuación*)

$F_{0.05,\nu_1,\nu_2}$

Grados de libertad del numerador (ν_1); Grados de libertad del denominador (ν_2)

ν_2 \ ν_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5	241.9	243.9	245.9	248.0	249.1	250.1	251.1	252.2	253.3	254.3
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.50
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.36
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73	1.67
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	1.64
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.55	1.43	1.35	1.25
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00

IV. Puntos porcentuales de la distribución F (*continuación*)

$F_{0.025,\nu_1,\nu_2}$

Grados de libertad del numerador (ν_1); Grados de libertad del denominador (ν_2)

ν_2 \ ν_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	647.8	799.5	864.2	899.6	921.8	937.1	948.2	956.7	963.3	968.6	976.7	984.9	993.1	997.2	1001	1006	1010	1014	1018
2	38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.37	39.39	39.40	39.41	39.43	39.45	39.46	39.46	39.47	39.48	39.49	39.50
3	17.44	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47	14.42	14.34	14.25	14.17	14.12	14.08	14.04	13.99	13.95	13.90
4	12.22	10.65	9.98	9.60	9.36	9.20	9.07	8.98	8.90	8.84	8.75	8.66	8.56	8.51	8.46	8.41	8.36	8.31	8.26
5	10.01	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68	6.62	6.52	6.43	6.33	6.28	6.23	6.18	6.12	6.07	6.02
6	8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52	5.46	5.37	5.27	5.17	5.12	5.07	5.01	4.96	4.90	4.85
7	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82	4.76	4.67	4.57	4.47	4.42	4.36	4.31	4.25	4.20	4.14
8	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36	4.30	4.20	4.10	4.00	3.95	3.89	3.84	3.78	3.73	3.67
9	7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03	3.96	3.87	3.77	3.67	3.61	3.56	3.51	3.45	3.39	3.33
10	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78	3.72	3.62	3.52	3.42	3.37	3.31	3.26	3.20	3.14	3.08
11	6.72	5.26	4.63	4.28	4.04	3.88	3.76	3.66	3.59	3.53	3.43	3.33	3.23	3.17	3.12	3.06	3.00	2.94	2.88
12	6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44	3.37	3.28	3.18	3.07	3.02	2.96	2.91	2.85	2.79	2.72
13	6.41	4.97	4.35	4.00	3.77	3.60	3.48	3.39	3.31	3.25	3.15	3.05	2.95	2.89	2.84	2.78	2.72	2.66	2.60
14	6.30	4.86	4.24	3.89	3.66	3.50	3.38	3.29	3.21	3.15	3.05	2.95	2.84	2.79	2.73	2.67	2.61	2.55	2.49
15	6.20	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.12	3.06	2.96	2.86	2.76	2.70	2.64	2.59	2.52	2.46	2.40
16	6.12	4.69	4.08	3.73	3.50	3.34	3.22	3.12	3.05	2.99	2.89	2.79	2.68	2.63	2.57	2.51	2.45	2.38	2.32
17	6.04	4.62	4.01	3.66	3.44	3.28	3.16	3.06	2.98	2.92	2.82	2.72	2.62	2.56	2.50	2.44	2.38	2.32	2.25
18	5.98	4.56	3.95	3.61	3.38	3.22	3.10	3.01	2.93	2.87	2.77	2.67	2.56	2.50	2.44	2.38	2.32	2.26	2.19
19	5.92	4.51	3.90	3.56	3.33	3.17	3.05	2.96	2.88	2.82	2.72	2.62	2.51	2.45	2.39	2.33	2.27	2.20	2.13
20	5.87	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84	2.77	2.68	2.57	2.46	2.41	2.35	2.29	2.22	2.16	2.09
21	5.83	4.42	3.82	3.48	3.25	3.09	2.97	2.87	2.80	2.73	2.64	2.53	2.42	2.37	2.31	2.25	2.18	2.11	2.04
22	5.79	4.38	3.78	3.44	3.22	3.05	2.93	2.84	2.76	2.70	2.60	2.50	2.39	2.33	2.27	2.21	2.14	2.08	2.00
23	5.75	4.35	3.75	3.41	3.18	3.02	2.90	2.81	2.73	2.67	2.57	2.47	2.36	2.30	2.24	2.18	2.11	2.04	1.97
24	5.72	4.32	3.72	3.38	3.15	2.99	2.87	2.78	2.70	2.64	2.54	2.44	2.33	2.27	2.21	2.15	2.08	2.01	1.94
25	5.69	4.29	3.69	3.35	3.13	2.97	2.85	2.75	2.68	2.61	2.51	2.41	2.30	2.24	2.18	2.12	2.05	1.98	1.91
26	5.66	4.27	3.67	3.33	3.10	2.94	2.82	2.73	2.65	2.59	2.49	2.39	2.28	2.22	2.16	2.09	2.03	1.95	1.88
27	5.63	4.24	3.65	3.31	3.08	2.92	2.80	2.71	2.63	2.57	2.47	2.36	2.25	2.19	2.13	2.07	2.00	1.93	1.85
28	5.61	4.22	3.63	3.29	3.06	2.90	2.78	2.69	2.61	2.55	2.45	2.34	2.23	2.17	2.11	2.05	1.98	1.91	1.83
29	5.59	4.20	1.61	3.27	3.04	2.88	2.76	2.67	2.59	2.53	2.43	2.32	2.21	2.15	2.09	2.03	1.96	1.89	1.81
30	5.57	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57	2.51	2.41	2.31	2.20	2.14	2.07	2.01	1.94	1.87	1.79
40	5.42	4.05	3.46	3.13	2.90	2.74	2.62	2.53	2.45	2.39	2.29	2.18	2.07	2.01	1.94	1.88	1.80	1.72	1.64
60	5.29	3.93	3.34	3.01	2.79	2.63	2.51	2.41	2.33	2.27	2.17	2.06	1.94	1.88	1.82	1.74	1.67	1.58	1.48
120	5.15	3.80	3.23	2.89	2.67	2.52	2.39	2.30	2.22	2.16	2.05	1.94	1.82	1.76	1.69	1.61	1.53	1.43	1.31
∞	5.02	3.69	3.12	2.79	2.57	2.41	2.29	2.19	2.11	2.05	1.94	1.83	1.71	1.64	1.57	1.48	1.39	1.27	1.00

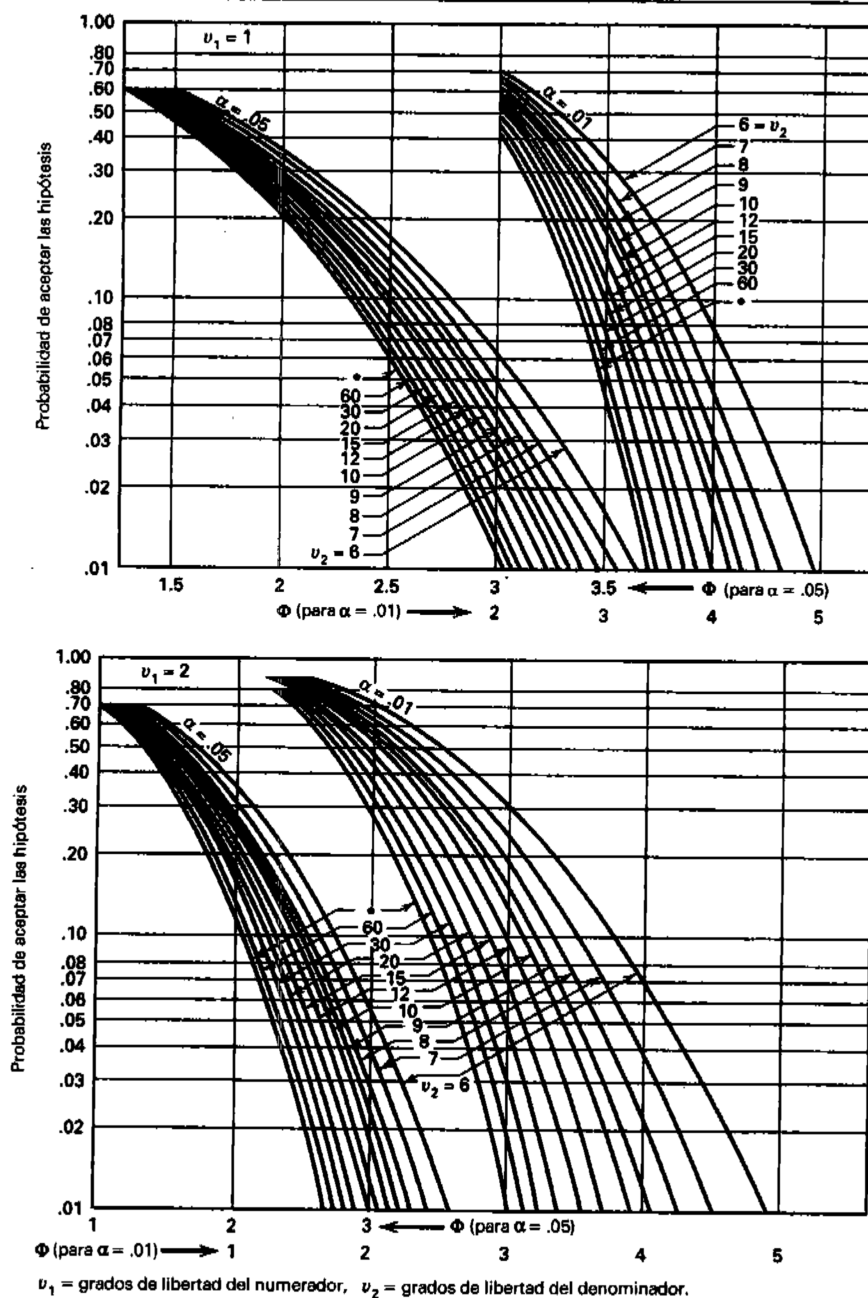
IV. Puntos porcentuales de la distribución F (*continuación*)

$F_{0.01,\nu_1,\nu_2}$

Grados de libertad del numerador (ν_1); Grados de libertad del denominador (ν_2)

ν_2 \ ν_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	4052	4999.5	5403	5625	5764	5859	5928	5982	6022	6056	6106	6157	6209	6235	6261	6287	6313	6339	6366
2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39	99.40	99.42	99.43	99.45	99.46	99.47	99.47	99.48	99.49	99.50
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35	27.23	27.05	26.87	26.69	26.00	26.50	26.41	26.32	26.22	26.13
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.55	14.37	14.20	14.02	13.93	13.84	13.75	13.65	13.56	13.46
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16	10.05	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72	7.56	7.40	7.31	7.23	7.14	7.06	6.97	6.88
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.74	5.65
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67	5.52	5.36	5.28	5.20	5.12	5.03	4.95	4.86
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.11	4.96	4.81	4.73	4.65	4.57	4.48	4.40	4.31
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00	3.91
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.40	4.25	4.10	4.02	3.94	3.86	3.78	3.69	3.60
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.45	3.36
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.25	3.17
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03	3.94	3.80	3.66	3.51	3.43	3.35	3.27	3.18	3.09	3.00
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96	2.87
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.84	2.75
17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.75	2.65
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51	3.37	3.23	3.08	3.00	2.92	2.84	2.75	2.66	2.57
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.30	3.15	3.00	2.92	2.84	2.76	2.67	2.58	2.49
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52	2.42
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31	3.17	3.03	2.88	2.80	2.72	2.64	2.55	2.46	2.36
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.40	2.31
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21	3.07	2.93	2.78	2.70	2.62	2.54	2.45	2.35	2.26
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31	2.21
25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22	3.13	2.99	2.85	2.70	2.62	2.54	2.45	2.36	2.27	2.17
26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18	3.09	2.96	2.81	2.66	2.58	2.50	2.42	2.33	2.23	2.13
27	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15	3.06	2.93	2.78	2.63	2.55	2.47	2.38	2.29	2.20	2.10
28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12	3.03	2.90	2.75	2.60	2.52	2.44	2.34	2.26	2.17	2.06
29	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09	3.00	2.87	2.73	2.57	2.49	2.41	2.33	2.23	2.14	2.03
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.11	2.01
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92	1.80
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.73	1.60
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.34	2.19	2.03	1.95	1.86	1.76	1.66	1.53	1.38
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.32	1.00

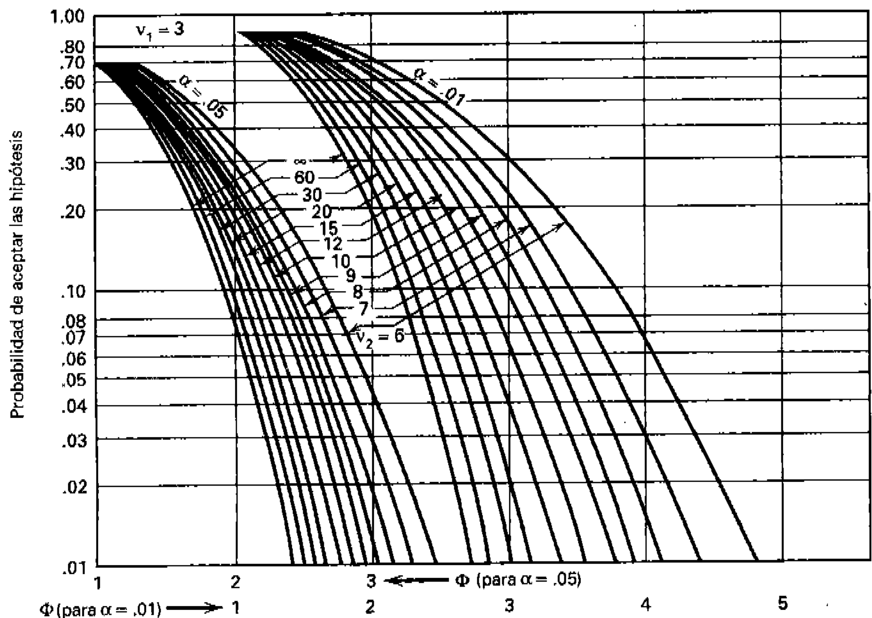
V. Curvas de operación característica para el análisis de varianza del modelo con efectos fijos[a]

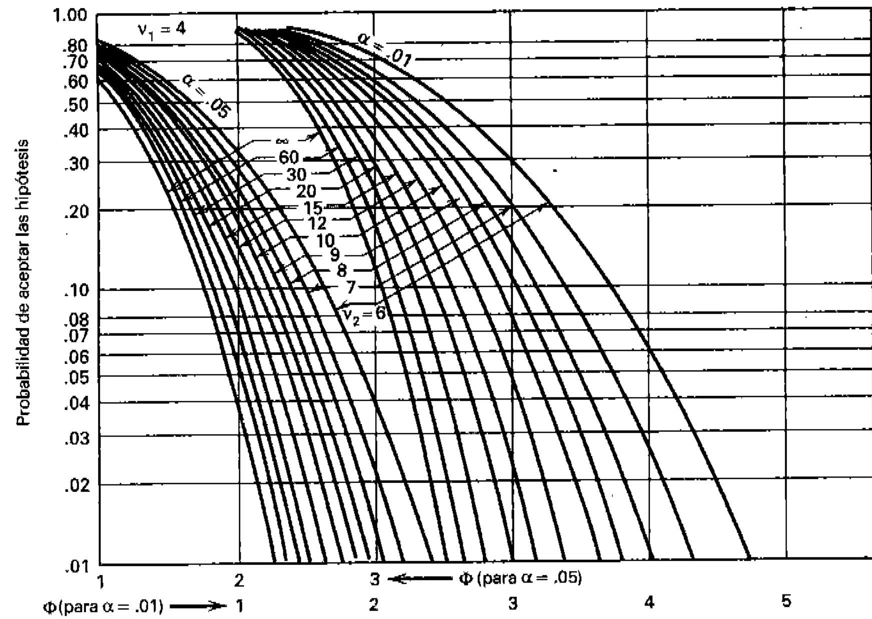


v_1 = grados de libertad del numerador, v_2 = grados de libertad del denominador.

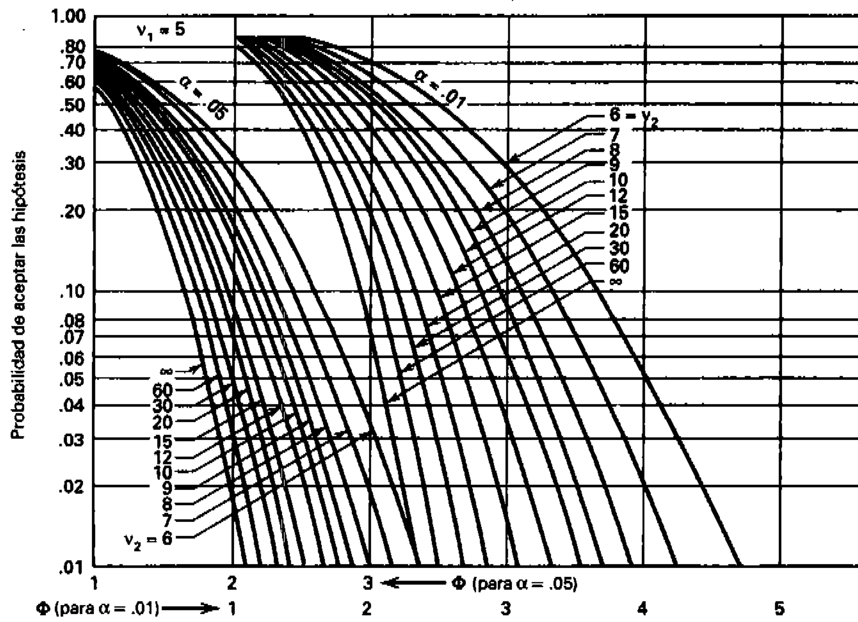
[a]Adaptada con permiso de *Biometrika Tables for Statisticians*, vol. 2, E.S. Pearson y H.O. Hartley, Cambridge University Press, Cambridge

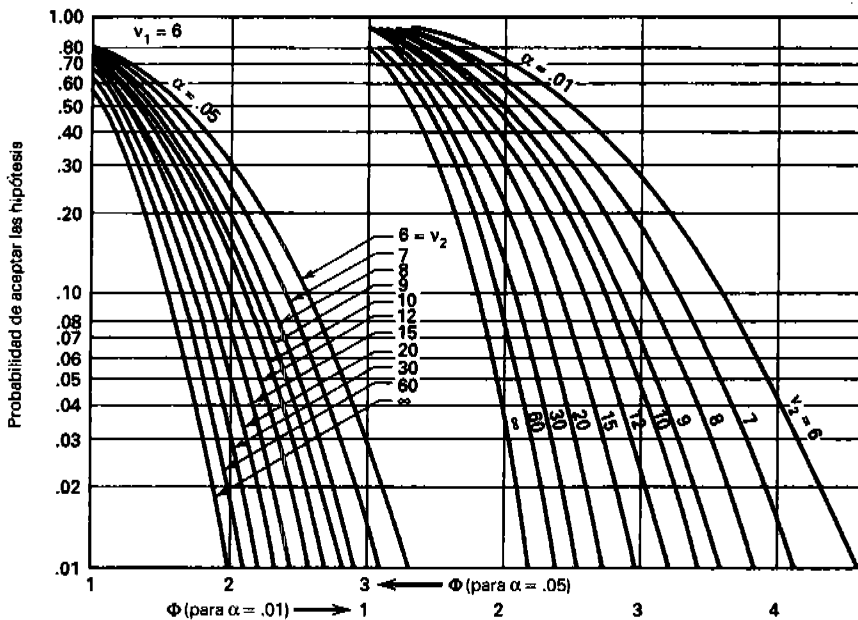
V. Curvas de operación característica para el análisis de varianza del modelo con efectos fijos (*continuación*)



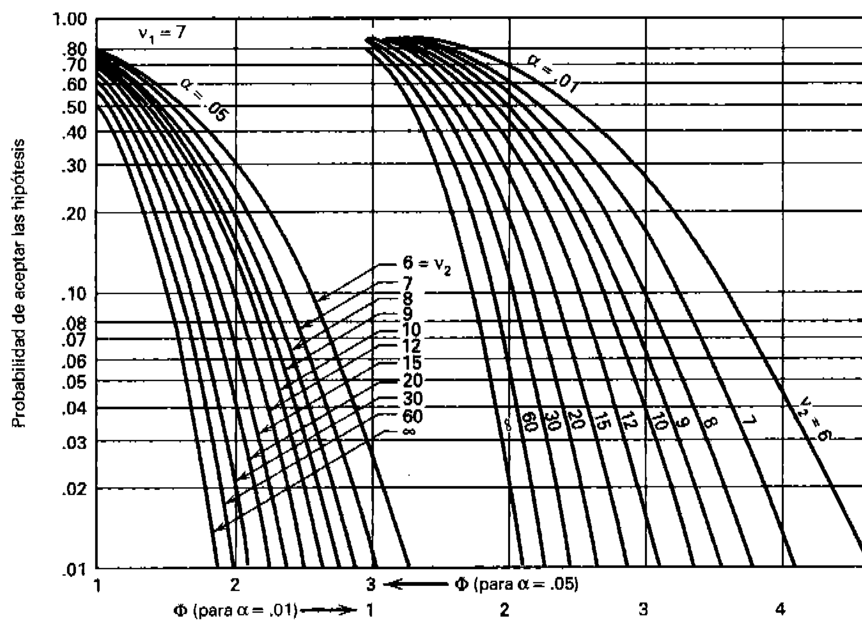


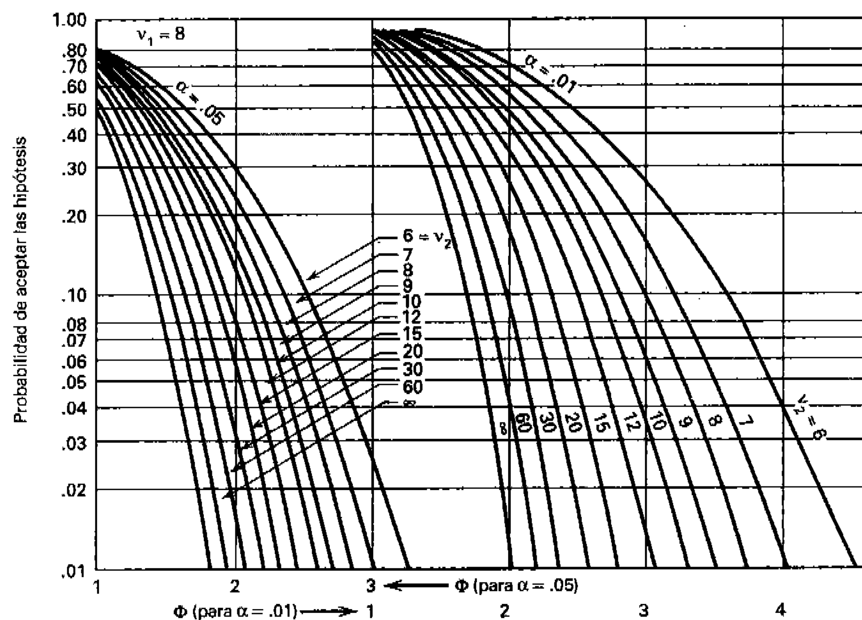
V. Curvas de operación característica para el análisis de varianza del modelo con efectos fijos (*continuación*)



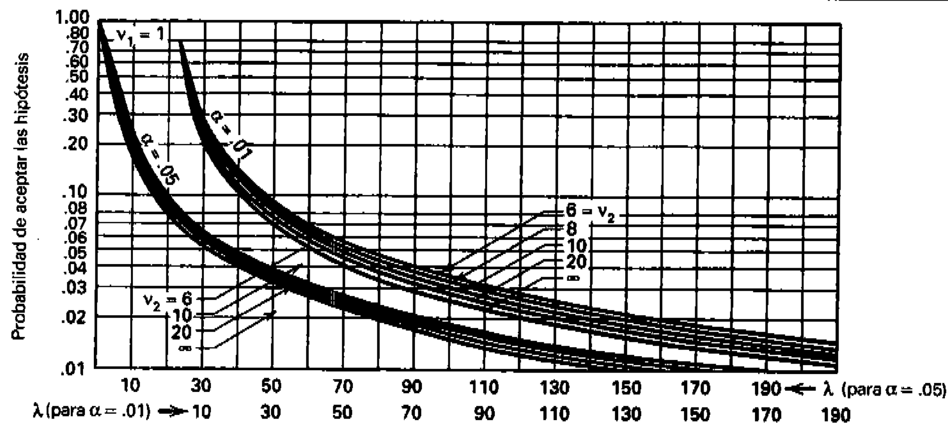


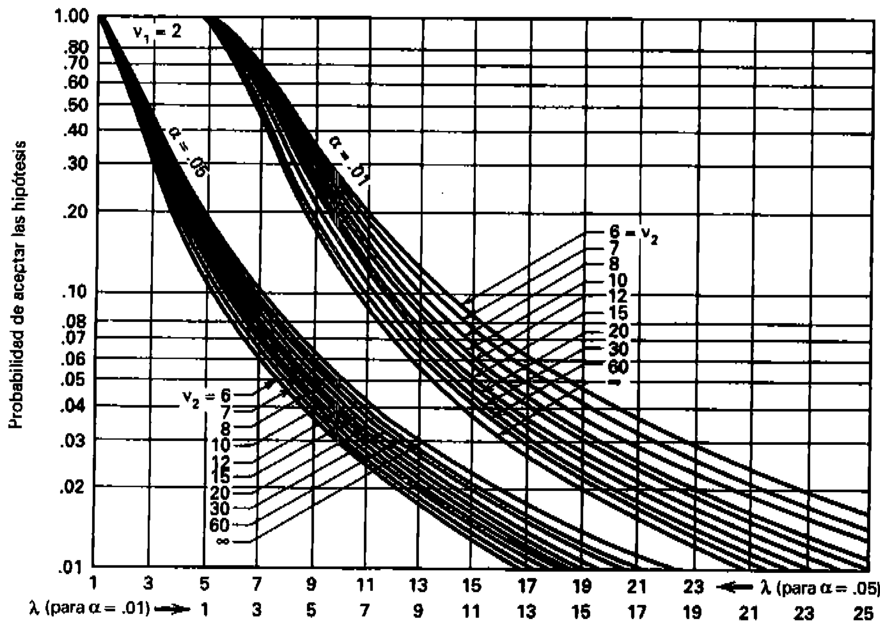
V. Curvas de operación característica para el análisis de varianza del modelo con efectos fijos (*continuación*)





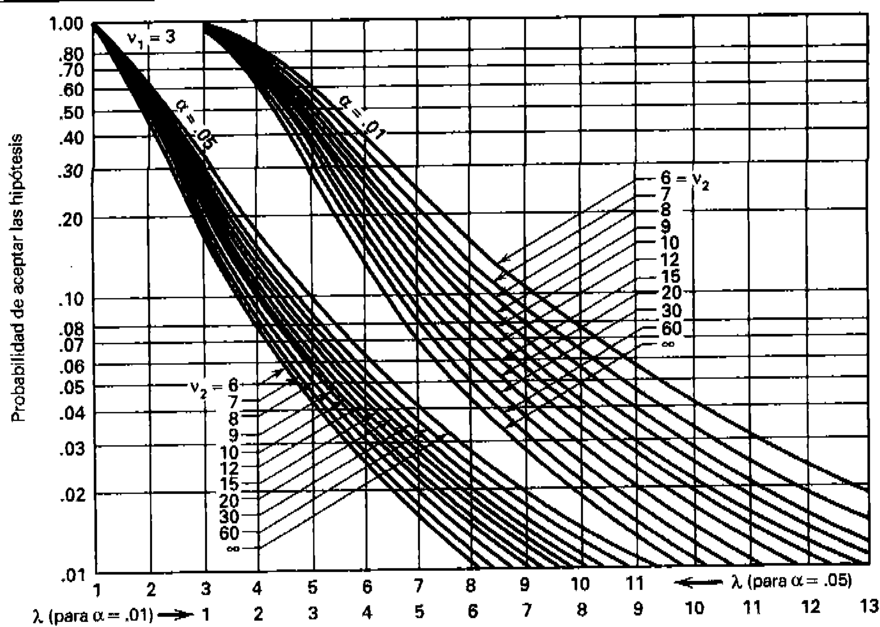
VI. Curvas de operación característica para el análisis de varianza del modelo con efectos aleatorios[a]

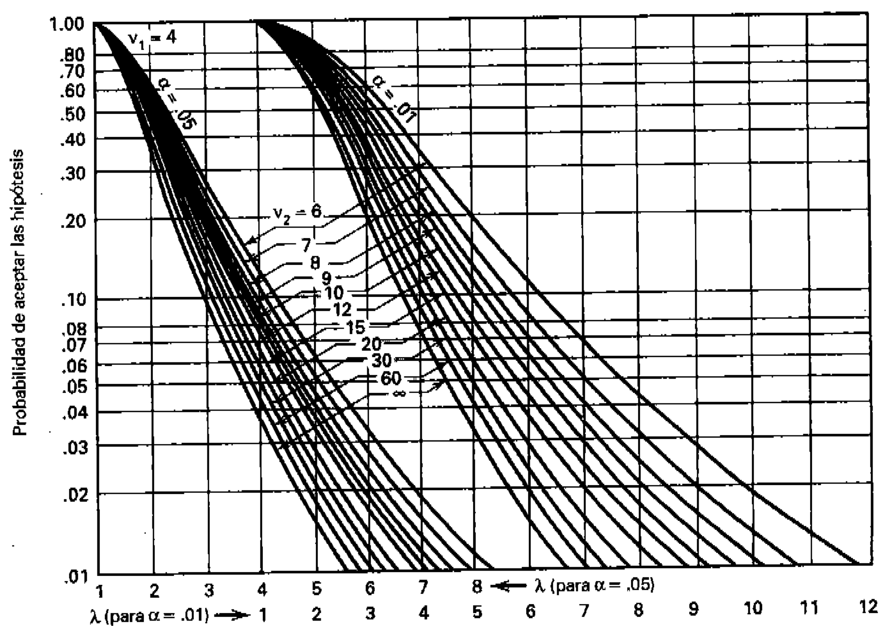




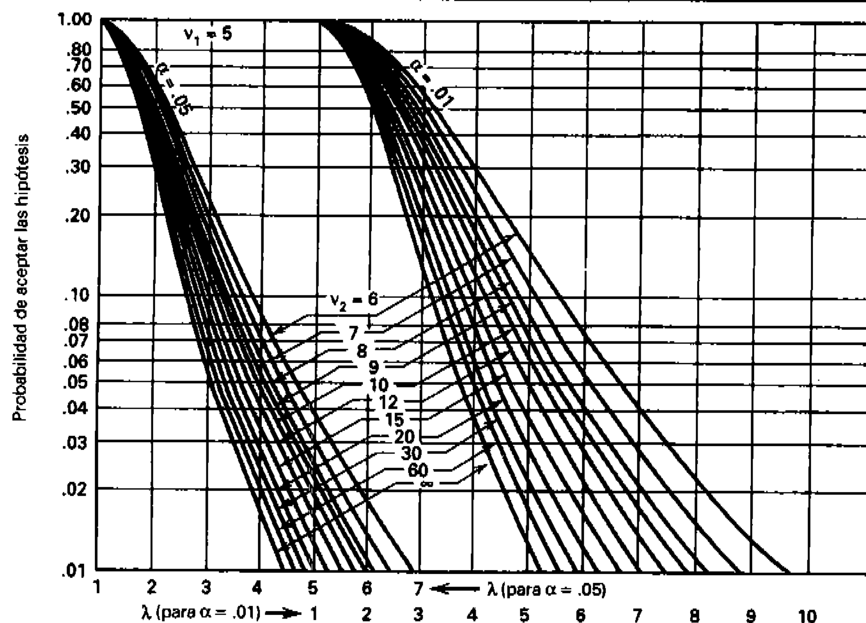
[a]Reproducida con permiso de *Engineering Statistics*, 2a. ed., A.H. Bowker y G.J. Lieberman, Prentice-Hall, Inc., Englewood Cliffs, NJ.

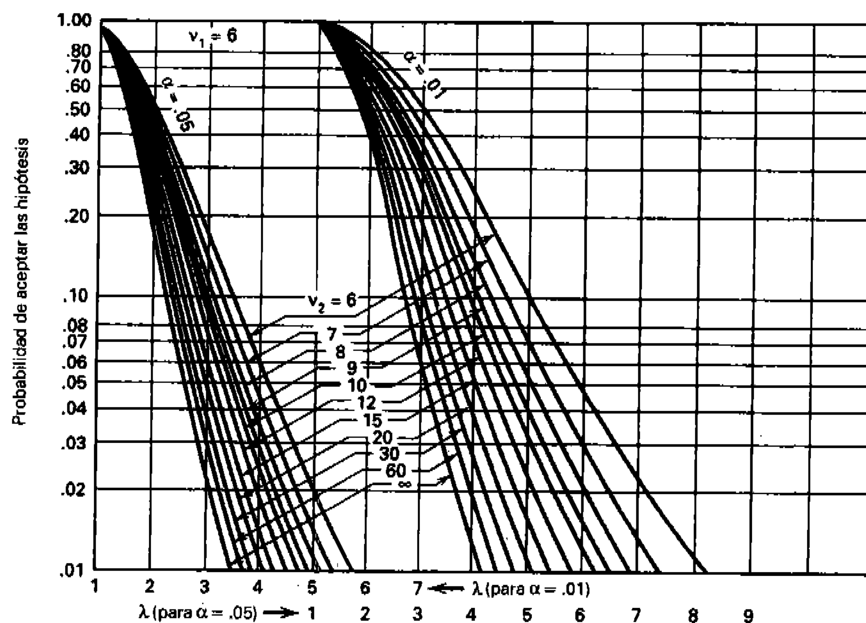
VI. Curvas de operación característica para el análisis de varianza del modelo con efectos aleatorios (*continuación*)



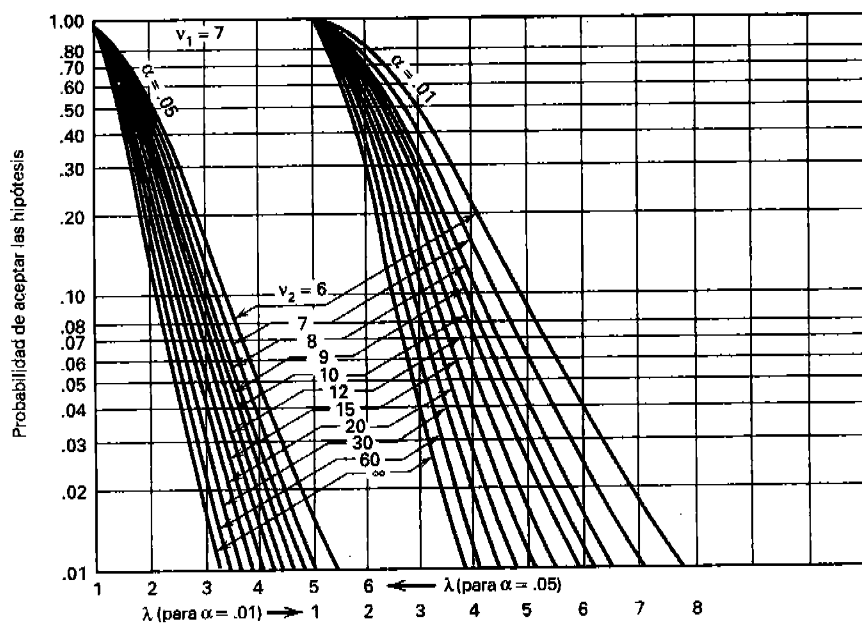


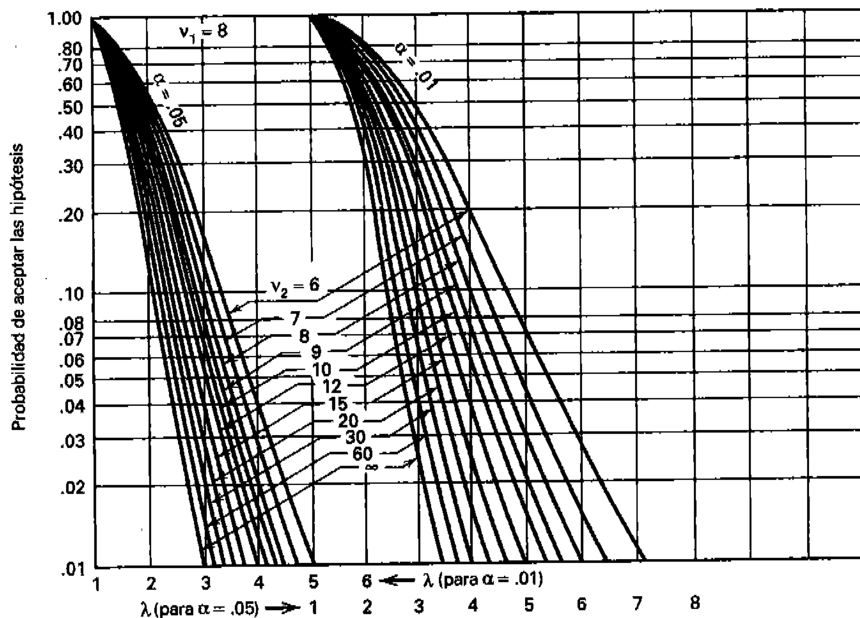
VI. Curvas de operación característica para el análisis de varianza del modelo con efectos aleatorios (*continuación*)





VI. Curvas de operación característica para el análisis de varianza del modelo con efectos aleatorios (*continuación*)





VII. Rangos significativos para la prueba del rango múltiple de Duncan[a]

$r_{0.01}(p, f)$

f	p = 2	3	4	5	6	7	8	9	10	20	50	100
1	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0
2	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0
3	8.26	8.5	8.6	8.7	8.8	8.9	8.9	9.0	9.0	9.3	9.3	9.3
4	6.51	6.8	6.9	7.0	7.1	7.1	7.2	7.2	7.3	7.5	7.5	7.5
5	5.70	5.96	6.11	6.18	6.26	6.33	6.40	6.44	6.5	6.8	6.8	6.8
6	5.24	5.51	5.65	5.73	5.81	5.88	5.95	6.00	6.0	6.3	6.3	6.3
7	4.95	5.22	5.37	5.45	5.53	5.61	5.69	5.73	5.8	6.0	6.0	6.0
8	4.74	5.00	5.14	5.23	5.32	5.40	5.47	5.51	5.5	5.8	5.8	5.8
9	4.60	4.86	4.99	5.08	5.17	5.25	5.32	5.36	5.4	5.7	5.7	5.7
10	4.48	4.73	4.88	4.96	5.06	5.13	5.20	5.24	5.28	5.55	5.55	5.55
11	4.39	4.63	4.77	4.86	4.94	5.01	5.06	5.12	5.15	5.39	5.39	5.39
12	4.32	4.55	4.68	4.76	4.84	4.92	4.96	5.02	5.07	5.26	5.26	5.26
13	4.26	4.48	4.62	4.69	4.74	4.84	4.88	4.94	4.98	5.15	5.15	5.15
14	4.21	4.42	4.55	4.63	4.70	4.78	4.83	4.87	4.91	5.07	5.07	5.07
15	4.17	4.37	4.50	4.58	4.64	4.72	4.77	4.81	4.84	5.00	5.00	5.00
16	4.13	4.34	4.45	4.54	4.60	4.67	4.72	4.76	4.79	4.94	4.94	4.94
17	4.10	4.30	4.41	4.50	4.56	4.63	4.68	4.73	4.75	4.89	4.89	4.89
18	4.07	4.27	4.38	4.46	4.53	4.59	4.64	4.68	4.71	4.85	4.85	4.85
19	4.05	4.24	4.35	4.43	4.50	4.56	4.61	4.64	4.67	4.82	4.82	4.82
20	4.02	4.22	4.33	4.40	4.47	4.53	4.58	4.61	4.65	4.79	4.79	4.79
30	3.89	4.06	4.16	4.22	4.32	4.36	4.41	4.45	4.48	4.65	4.71	4.71
40	3.82	3.99	4.10	4.17	4.24	4.30	4.34	4.37	4.41	4.59	4.69	4.69
60	3.76	3.92	4.03	4.12	4.17	4.23	4.27	4.31	4.34	4.53	4.66	4.66
100	3.71	3.86	3.98	4.06	4.11	4.17	4.21	4.25	4.29	4.48	4.64	4.65
∞	3.64	3.80	3.90	3.98	4.04	4.09	4.14	4.17	4.20	4.41	4.60	4.68

f = grados de libertad

Reproducida con permiso de "Multiple Range and Multiple F Tests", D.B. Duncan, *Biometrics*, vol. 1, no. 1, pp. 1-42

$r_{0.05}(p, f)$

f	p = 2	3	4	5	6	7	8	9	10	20	50	100
1	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0
2	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09
3	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50
4	3.93	4.01	4.02	4.02	4.02	4.02	4.02	4.02	4.02	4.02	4.02	4.02
5	3.64	3.74	3.79	3.83	3.83	3.83	3.83	3.83	3.83	3.83	3.83	3.83
6	3.46	3.58	3.64	3.68	3.68	3.68	3.68	3.68	3.68	3.68	3.68	3.68
7	3.35	3.47	3.54	3.58	3.60	3.61	3.61	3.61	3.61	3.61	3.61	3.61
8	3.26	3.39	3.47	3.52	3.55	3.56	3.56	3.56	3.56	3.56	3.56	3.56
9	3.20	3.34	3.41	3.47	3.50	3.52	3.52	3.52	3.52	3.52	3.52	3.52
10	3.15	3.30	3.37	3.43	3.46	3.47	3.47	3.47	3.47	3.48	3.48	3.48
11	3.11	3.27	3.35	3.39	3.43	3.44	3.45	3.46	3.46	3.48	3.48	3.48
12	3.08	3.23	3.33	3.36	3.40	3.42	3.44	3.44	3.46	3.48	3.48	3.48
13	3.06	3.21	3.30	3.35	3.38	3.41	3.42	3.44	3.45	3.47	3.47	3.47
14	3.03	3.18	3.27	3.33	3.37	3.39	3.41	3.42	3.44	3.47	3.47	3.47
15	3.01	3.16	3.25	3.31	3.36	3.38	3.40	3.42	3.43	3.47	3.47	3.47
16	3.00	3.15	3.23	3.30	3.34	3.37	3.39	3.41	3.43	3.47	3.47	3.47
17	2.98	3.13	3.22	3.28	3.33	3.36	3.38	3.40	3.42	3.47	3.47	3.47
18	2.97	3.12	3.21	3.27	3.32	3.35	3.37	3.39	3.41	3.47	3.47	3.47
19	2.96	3.11	3.19	3.26	3.31	3.35	3.37	3.39	3.41	3.47	3.47	3.47
20	2.95	3.10	3.18	3.25	3.30	3.34	3.36	3.38	3.40	3.47	3.47	3.47
30	2.89	3.04	3.12	3.20	3.25	3.29	3.32	3.35	3.37	3.47	3.47	3.47
40	2.86	3.01	3.10	3.17	3.22	3.27	3.30	3.33	3.35	3.47	3.47	3.47
60	2.83	2.98	3.08	3.14	3.20	3.24	3.28	3.31	3.33	3.47	3.48	3.48
100	2.80	2.95	3.05	3.12	3.18	3.22	3.26	3.29	3.32	3.47	3.53	3.53
∞	2.77	2.92	3.02	3.09	3.15	3.19	3.23	3.26	3.29	3.47	3.61	3.67

VIII. Puntos porcentuales del estadístico del rango studentizado[a]
$q_{0.01}(p, f)$

	p																		
f	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	90.0	135	164	186	202	216	227	237	246	253	260	266	272	272	282	286	290	294	298
2	14.0	19.0	22.3	24.7	26.6	28.2	29.5	30.7	31.7	32.6	33.4	31.4	34.8	35.4	36.0	36.5	37.0	37.5	37.9
3	8.26	10.6	12.2	13.3	14.2	15.0	15.6	16.2	16.7	17.1	17.5	17.9	18.2	18.5	18.8	19.1	19.3	19.5	19.8
4	6.51	8.12	9.17	9.96	10.6	11.1	11.5	11.9	12.3	12.6	12.8	13.1	13.3	13.5	13.7	13.9	14.1	14.2	14.4
5	5.70	6.97	7.80	8.42	8.91	9.32	9.67	9.97	10.24	10.48	10.70	10.89	11.08	11.24	11.40	11.55	11.68	11.81	11.93
6	5.24	6.33	7.03	7.56	7.97	8.32	8.61	8.87	9.10	9.30	9.49	9.65	9.81	9.95	10.08	10.21	10.32	10.43	10.54
7	4.95	5.92	6.54	7.01	7.37	7.68	7.94	8.17	8.37	8.55	8.71	8.86	9.00	9.12	9.24	9.35	9.46	9.55	9.65
8	4.74	5.63	6.20	6.63	6.96	7.24	7.47	7.68	7.87	8.03	8.18	8.31	8.44	8.55	8.66	8.76	8.85	8.94	9.03
9	4.60	5.43	5.96	6.35	6.66	6.91	7.13	7.32	7.49	7.65	7.78	7.91	8.03	8.13	8.23	8.32	8.41	8.49	8.57
10	4.48	5.27	5.77	6.14	6.43	6.67	6.87	7.05	7.21	7.36	7.48	7.60	7.71	7.81	7.91	7.99	8.07	8.15	8.22
11	4.39	5.14	5.62	5.97	6.25	6.48	6.67	6.84	6.99	7.13	7.25	7.36	7.46	7.56	7.65	7.73	7.81	7.88	7.95
12	4.32	5.04	5.50	5.84	6.10	6.32	6.51	6.67	6.81	6.94	7.06	7.17	7.26	7.36	7.44	7.52	7.59	7.66	7.73
13	4.26	4.96	5.40	5.73	5.98	6.19	6.37	6.53	6.67	6.79	6.90	7.01	7.10	7.19	7.27	7.34	7.42	7.48	7.55
14	4.21	4.89	5.32	5.63	5.88	6.08	6.26	6.41	6.54	6.66	6.77	6.87	6.96	7.05	7.12	7.20	7.27	7.33	7.39
15	4.17	4.83	5.25	5.56	5.80	5.99	6.16	6.31	6.44	6.55	6.66	6.76	6.84	6.93	7.00	7.07	7.14	7.20	7.26
16	4.13	4.78	5.19	5.49	5.72	5.92	6.08	6.22	6.35	6.46	6.56	6.66	6.74	6.82	6.90	6.97	7.03	7.09	7.15
17	4.10	4.74	5.14	5.43	5.66	5.85	6.01	6.15	6.27	6.38	6.48	6.57	6.66	6.73	6.80	6.87	6.94	7.00	7.05
18	4.07	4.70	5.09	5.38	5.60	5.79	5.94	6.08	6.20	6.31	6.41	6.50	6.58	6.65	6.72	6.79	6.85	6.91	6.96
19	4.05	4.67	5.05	5.33	5.55	5.73	5.89	6.02	6.14	6.25	6.34	6.43	6.51	6.58	6.65	6.72	6.78	6.84	6.89
20	4.02	4.64	5.02	5.29	5.51	5.69	5.84	5.97	6.09	6.19	6.29	6.37	6.45	6.52	6.59	6.65	6.71	6.76	6.82
24	3.96	4.54	4.91	5.17	5.37	5.54	5.69	5.81	5.92	6.02	6.11	6.19	6.26	6.33	6.39	6.45	6.51	6.56	6.61
30	3.89	4.45	4.80	5.05	5.24	5.40	5.54	5.65	5.76	5.85	5.93	6.01	6.08	6.14	6.20	6.26	6.31	6.36	6.41
40	3.82	4.37	4.70	4.93	5.11	5.27	5.39	5.50	5.60	5.69	5.77	5.84	5.90	5.96	6.02	6.07	6.12	6.17	6.21
60	3.76	4.28	4.60	4.82	4.99	5.13	5.25	5.36	5.45	5.53	5.60	5.67	5.73	5.79	5.84	5.89	5.93	5.98	6.02
120	3.70	4.20	4.50	4.71	4.87	5.01	5.12	5.21	5.30	5.38	5.44	5.51	5.56	5.61	5.66	5.71	5.75	5.79	5.83
∞	3.64	4.12	4.40	4.60	4.76	4.88	4.99	5.08	5.16	5.23	5.29	5.35	5.40	5.45	5.49	5.54	5.57	5.61	5.65

f = grados de libertad

[a] De J.M. May, "Extended and Corrected Tables of the Upper Percentage Points of the Studentized Range", *Biometrika*, vol. 39, pp. 192-193. Reproducida con permiso de los fideicomisarios de *Biometrika*

VIII. Puntos porcentuales del estadístico del rango studentizado (*continuación*)

$q_{0.05}(p, f)$

f	p = 2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	18.1	26.7	32.8	37.2	40.5	43.1	45.4	47.3	49.1	50.6	51.9	53.2	54.3	55.4	56.3	57.2	58.0	58.8	59.6
2	6.09	8.28	9.80	10.89	11.73	12.43	13.03	13.54	13.99	14.39	14.75	15.08	15.38	15.65	15.91	16.14	16.36	16.57	16.77
3	4.50	5.88	6.83	7.51	8.04	8.47	8.85	9.18	9.46	9.72	9.95	10.16	10.35	10.52	10.69	10.84	10.98	11.12	11.24
4	3.93	5.00	5.76	6.31	6.73	7.06	7.35	7.60	7.83	8.03	8.21	8.37	8.52	8.67	8.80	8.92	9.03	9.14	9.24
5	3.64	4.60	5.22	5.67	6.03	6.33	6.58	6.80	6.99	7.17	7.32	7.47	7.60	7.72	7.83	7.93	8.03	8.12	8.21
6	3.46	4.34	4.90	5.31	5.63	5.89	6.12	6.32	6.49	6.65	6.79	6.92	7.04	7.14	7.24	7.34	7.43	7.51	7.59
7	3.34	4.16	4.68	5.06	5.35	5.59	5.80	5.99	6.15	6.29	6.42	6.54	6.65	6.75	6.84	6.93	7.01	7.08	7.16
8	3.26	4.04	4.53	4.89	5.17	5.40	5.60	5.77	5.92	6.05	6.18	6.29	6.39	6.48	6.57	6.65	6.73	6.80	6.87
9	3.20	3.95	4.42	4.76	5.02	5.24	5.43	5.60	5.74	5.87	5.98	6.09	6.19	6.28	6.36	6.44	6.51	6.58	6.65
10	3.15	3.88	4.33	4.66	4.91	5.12	5.30	5.46	5.60	5.72	5.83	5.93	6.03	6.12	6.20	6.27	6.34	6.41	6.47
11	3.11	3.82	4.26	4.58	4.82	5.03	5.20	5.35	5.49	5.61	5.71	5.81	5.90	5.98	6.06	6.14	6.20	6.27	6.33
12	3.08	3.77	4.20	4.51	4.75	4.95	5.12	5.27	5.40	5.51	5.61	5.71	5.80	5.88	5.95	6.02	6.09	6.15	6.21
13	3.06	3.73	4.15	4.46	4.69	4.88	5.05	5.19	5.32	5.43	5.53	5.63	5.71	5.79	5.86	5.93	6.00	6.06	6.11
14	3.03	3.70	4.11	4.41	4.64	4.83	4.99	5.13	5.25	5.36	5.46	5.56	5.64	5.72	5.79	5.86	5.92	5.98	6.03
15	3.01	3.67	4.08	4.37	4.59	4.78	4.94	5.08	5.20	5.31	5.40	5.49	5.57	5.65	5.72	5.79	5.85	5.91	5.96
16	3.00	3.65	4.05	4.34	4.56	4.74	4.90	5.03	5.15	5.26	5.35	5.44	5.52	5.59	5.66	5.73	5.79	5.84	5.90
17	2.98	3.62	4.02	4.31	4.52	4.70	4.86	4.99	5.11	5.21	5.31	5.39	5.47	5.55	5.61	5.68	5.74	5.79	5.84
18	2.97	3.61	4.00	4.28	4.49	4.67	4.83	4.96	5.07	5.17	5.27	5.35	5.43	5.50	5.57	5.63	5.69	5.74	5.79
19	2.96	3.59	3.98	4.26	4.47	4.64	4.79	4.92	5.04	5.14	5.23	5.32	5.39	5.46	5.53	5.59	5.65	5.70	5.75
20	2.95	3.58	3.96	4.24	4.45	4.62	4.77	4.90	5.01	5.11	5.20	5.28	5.36	5.43	5.50	5.56	5.61	5.66	5.71
24	2.92	3.53	3.90	4.17	4.37	4.54	4.68	4.81	4.92	5.01	5.10	5.18	5.25	5.32	5.38	5.44	5.50	5.55	5.59
30	2.89	3.48	3.84	4.11	4.30	4.46	4.60	4.72	4.83	4.92	5.00	5.08	5.15	5.21	5.27	5.33	5.38	5.43	5.48
40	2.86	3.44	3.79	4.04	4.23	4.39	4.52	4.63	4.74	4.82	4.90	4.98	5.05	5.11	5.17	5.22	5.27	5.32	5.36
60	2.83	3.40	3.74	3.98	4.16	4.31	4.44	4.55	4.65	4.73	4.81	4.88	4.94	5.00	5.06	5.11	5.15	5.20	5.24
120	2.80	3.36	3.69	3.92	4.10	4.24	4.36	4.47	4.56	4.64	4.71	4.78	4.84	4.90	4.95	5.00	5.04	5.09	5.13
∞	2.77	3.32	3.63	3.86	4.03	4.17	4.29	4.39	4.47	4.55	4.62	4.68	4.74	4.80	4.84	4.98	4.93	4.97	5.01

IX. Valores críticos para la prueba de Dunnett para comparar tratamientos con un control[a]

$d_{0.05}(a-1, f)$

Comparaciones de dos colas

	$a-1$ = número de medias de tratamientos (sin incluir el control)								
f	1	2	3	4	5	6	7	8	9
5	2.57	3.03	3.29	3.48	3.62	3.73	3.82	3.90	3.97
6	2.45	2.86	3.10	3.26	3.39	3.49	3.57	3.64	3.71
7	2.36	2.75	2.97	3.12	3.24	3.33	3.41	3.47	3.53
8	2.31	2.67	2.88	3.02	3.13	3.22	3.29	3.35	3.41
9	2.26	2.61	2.81	2.95	3.05	3.14	3.20	3.26	3.32
10	2.23	2.57	2.76	2.89	2.99	3.07	3.14	3.19	3.24
11	2.20	2.53	2.72	2.84	2.94	3.02	3.08	3.14	3.19
12	2.18	2.50	2.68	2.81	2.90	2.98	3.04	3.09	3.14
13	2.16	2.48	2.65	2.78	2.87	2.94	3.00	3.06	3.10
14	2.14	2.46	2.63	2.75	2.84	2.91	2.97	3.02	3.07
15	2.13	2.44	2.61	2.73	2.82	2.89	2.95	3.00	3.04
16	2.12	2.42	2.59	2.71	2.80	2.87	2.92	2.97	3.02
17	2.11	2.41	2.58	2.69	2.78	2.85	2.90	2.95	3.00
18	2.10	2.40	2.56	2.68	2.76	2.83	2.89	2.94	2.98
19	2.09	2.39	2.55	2.66	2.75	2.81	2.87	2.92	2.96
20	2.09	2.38	2.54	2.65	2.73	2.80	2.86	2.90	2.95
24	2.06	2.35	2.51	2.61	2.70	2.76	2.81	2.86	2.90
30	2.04	2.32	2.47	2.58	2.66	2.72	2.77	2.82	2.86
40	2.02	2.29	2.44	2.54	2.62	2.68	2.73	2.77	2.81
60	2.00	2.27	2.41	2.51	2.58	2.64	2.69	2.73	2.77
120	1.98	2.24	2.38	2.47	2.55	2.60	2.65	2.69	2.73
∞	1.96	2.21	2.35	2.44	2.51	2.57	2.61	2.65	2.69

f = grados de libertad

[a]Reproducida con permiso de C.W. Dunnett, "New Tables for Multiple Comparison with a Control", *Biometrics*, vol. 20, no. 3, y de C.W. Dunnett, "A Multiple Comparison Procedure for Comparing Several Treatments with a Control", *Journal of the American Statistical Association*, vol. 50

IX. Valores críticos para la prueba de Dunnett para comparar tratamientos con un control[a]

$d_{0.01}(a-1, f)$

Comparaciones de dos colas (*continuación*)

	$a-1$ = número de medias de tratamientos (sin incluir el control)								
f	1	2	3	4	5	6	7	8	9
5	4.03	4.63	4.98	5.22	5.41	5.56	5.69	5.80	5.89
6	3.71	4.21	4.51	4.71	4.87	5.00	5.10	5.20	5.28
7	3.50	3.95	4.21	4.39	4.53	4.64	4.74	4.82	4.89
8	3.36	3.77	4.00	4.17	4.29	4.40	4.48	4.56	4.62
9	3.25	3.63	3.85	4.01	4.12	4.22	4.30	4.37	4.43
10	3.17	3.53	3.74	3.88	3.99	4.08	4.16	4.22	4.28
11	3.11	3.45	3.65	3.79	3.89	3.98	4.05	4.11	4.16
12	3.05	3.39	3.58	3.71	3.81	3.89	3.96	4.02	4.07
13	3.01	3.33	3.52	3.65	3.74	3.82	3.89	3.94	3.99
14	2.98	3.29	3.47	3.59	3.69	3.76	3.83	3.88	3.93
15	2.95	3.25	3.43	3.55	3.64	3.71	3.78	3.83	3.88
16	2.92	3.22	3.39	3.51	3.60	3.67	3.73	3.78	3.83
17	2.90	3.19	3.36	3.47	3.56	3.63	3.69	3.74	3.79
18	2.88	3.17	3.33	3.44	3.53	3.60	3.66	3.71	3.75
19	2.86	3.15	3.31	3.42	3.50	3.57	3.63	3.68	3.72
20	2.85	3.13	3.29	3.40	3.48	3.55	3.60	3.65	3.69
24	2.80	3.07	3.22	3.32	3.40	3.47	3.52	3.57	3.61
30	2.75	3.01	3.15	3.25	3.33	3.39	3.44	3.49	3.52
40	2.70	2.95	3.09	3.19	3.26	3.32	3.37	3.41	3.44
60	2.66	2.90	3.03	3.12	3.19	3.25	3.29	3.33	3.37
120	2.62	2.85	2.97	3.06	3.12	3.18	3.22	3.26	3.29
∞	2.58	2.79	2.92	3.00	3.06	3.11	3.15	3.19	3.22

$d_{0.05}(a-1, f)$

Comparaciones de una cola

	$a-1$ = número de medias de tratamientos (sin incluir el control)								
f	1	2	3	4	5	6	7	8	9
5	2.02	2.44	2.68	2.85	2.98	3.08	3.16	3.24	3.30
6	1.94	2.34	2.56	2.71	2.83	2.92	3.00	3.07	3.12
7	1.89	2.27	2.48	2.62	2.73	2.82	2.89	2.95	3.01
8	1.86	2.22	2.42	2.55	2.66	2.74	2.81	2.87	2.92
9	1.83	2.18	2.37	2.50	2.60	2.68	2.75	2.81	2.86
10	1.81	2.15	2.34	2.47	2.56	2.64	2.70	2.76	2.81
11	1.80	2.13	2.31	2.44	2.53	2.60	2.67	2.72	2.77
12	1.78	2.11	2.29	2.41	2.50	2.58	2.64	2.69	2.74
13	1.77	2.09	2.27	2.39	2.48	2.55	2.61	2.66	2.71
14	1.76	2.08	2.25	2.37	2.46	2.53	2.59	2.64	2.69
15	1.75	2.07	2.24	2.36	2.44	2.51	2.57	2.62	2.67
16	1.75	2.06	2.23	2.34	2.43	2.50	2.56	2.61	2.65
17	1.74	2.05	2.22	2.33	2.42	2.49	2.54	2.59	2.64
18	1.73	2.04	2.21	2.32	2.41	2.48	2.53	2.58	2.62
19	1.73	2.03	2.20	2.31	2.40	2.47	2.52	2.57	2.61
20	1.72	2.03	2.19	2.30	2.39	2.46	2.51	2.56	2.60
24	1.71	2.01	2.17	2.28	2.36	2.43	2.48	2.53	2.57
30	1.70	1.99	2.15	2.25	2.33	2.40	2.45	2.50	2.54
40	1.68	1.97	2.13	2.23	2.31	2.37	2.42	2.47	2.51
60	1.67	1.95	2.10	2.21	2.28	2.35	2.39	2.44	2.48
120	1.66	1.93	2.08	2.18	2.26	2.32	2.37	2.41	2.45
∞	1.64	1.92	2.06	2.16	2.23	2.29	2.34	2.38	2.42

IX. Valores críticos para la prueba de Dunnett para comparar tratamientos con un control
$d_{0.01}(a-1, f)$
Comparaciones de una cola (*continuación*)

	$a-1$ = número de medias de tratamientos (sin incluir el control)								
f	1	2	3	4	5	6	7	8	9
5	3.37	3.90	4.21	4.43	4.60	4.73	4.85	4.94	5.03
6	3.14	3.61	3.88	4.07	4.21	4.33	4.43	4.51	4.59
7	3.00	3.42	3.66	3.83	3.96	4.07	4.15	4.23	4.30
8	2.90	3.29	3.51	3.67	3.79	3.88	3.96	4.03	4.09
9	2.82	3.19	3.40	3.55	3.66	3.75	3.82	3.89	3.94
10	2.76	3.11	3.31	3.45	3.56	3.64	3.71	3.78	3.83
11	2.72	3.06	3.25	3.38	3.48	3.56	3.63	3.69	3.74
12	2.68	3.01	3.19	3.32	3.42	3.50	3.56	3.62	3.67
13	2.65	2.97	3.15	3.27	3.37	3.44	3.51	3.56	3.61
14	2.62	2.94	3.11	3.23	3.32	3.40	3.46	3.51	3.56
15	2.60	2.91	3.08	3.20	3.29	3.36	3.42	3.47	3.52
16	2.58	2.88	3.05	3.17	3.26	3.33	3.39	3.44	3.48
17	2.57	2.86	3.03	3.14	3.23	3.30	3.36	3.41	3.45
18	2.55	2.84	3.01	3.12	3.21	3.27	3.33	3.38	3.42
19	2.54	2.83	2.99	3.10	3.18	3.25	3.31	3.36	3.40
20	2.53	2.81	2.97	3.08	3.17	3.23	3.29	3.34	3.38
24	2.49	2.77	2.92	3.03	3.11	3.17	3.22	3.27	3.31
30	2.46	2.72	2.87	2.97	3.05	3.11	3.16	3.21	3.24
40	2.42	2.68	2.82	2.92	2.99	3.05	3.10	3.14	3.18
60	2.39	2.64	2.78	2.87	2.94	3.00	3.04	3.08	3.12
120	2.36	2.60	2.73	2.82	2.89	2.94	2.99	3.03	3.06
∞	2.33	2.56	2.68	2.77	2.84	2.89	2.93	2.97	3.00

X. Coeficientes de polinomios ortogonales[a]

		$n=3$		$n=4$			$n=5$				$n=6$					$n=7$					
	X_j	P_1	P_2	P_1	P_2	P_3	P_1	P_2	P_3	P_4	P_1	P_2	P_3	P_4	P_5	P_1	P_2	P_3	P_4	P_5	P_6
	1	−1	1	−3	1	−1	−2	2	−1	1	−5	5	−5	1	−1	−3	5	−1	3	−1	1
	2	0	−2	−1	−1	3	1	−1	2	−4	−3	−1	7	−3	5	−2	0	1	−7	4	−6
	3	1	1	1	−1	−3	0	−2	0	6	−1	−4	4	2	−10	−1	−3	1	1	−5	15
	4			3	1	1	1	−1	−2	−4	1	−4	−4	2	10	0	−4	0	6	0	−20
	5						2	2	1	1	3	−1	−7	−3	−5	1	−3	−1	1	5	15
	6										5	5	5	1	1	2	0	−1	−7	−4	−6
	7															3	5	1	3	1	1
$\sum_{j=1}^{n} \{P_i(X_j)\}^2$		2	6	20	4	20	10	14	10	70	70	84	180	28	252	28	84	6	154	84	924
	λ	1	3	2	1	$\frac{10}{3}$	1	1	$\frac{5}{6}$	$\frac{35}{12}$	2	$\frac{3}{2}$	$\frac{5}{3}$	$\frac{7}{12}$	$\frac{21}{10}$	1	1	$\frac{1}{6}$	$\frac{7}{12}$	$\frac{7}{20}$	$\frac{77}{60}$

		$n=8$						$n=9$						$n=10$					
	X_j	P_1	P_2	P_3	P_4	P_5	P_6	P_1	P_2	P_3	P_4	P_5	P_6	P_1	P_2	P_3	P_4	P_5	P_6
	1	−7	7	−7	7	−7	1	−4	28	−14	14	−4	4	−9	6	−42	18	−6	3
	2	−5	1	5	−13	23	−5	−3	7	7	−21	11	−17	−7	2	14	−22	14	−11
	3	−3	−3	7	−3	−17	9	−2	−8	13	−11	−4	22	−5	−1	35	−17	−1	10
	4	−1	−5	3	9	−15	−5	−1	−17	9	9	−9	1	−3	−3	31	3	−11	6
	5	1	−5	−3	9	15	−5	0	−20	0	18	0	−20	−1	−4	12	18	−6	−8
	6	3	−3	−7	−3	17	9	1	−17	−9	9	9	1	1	−4	−12	18	6	−8
	7	5	1	−5	−13	−23	−5	2	−8	−13	−11	4	22	3	−3	−31	3	11	6
	8	7	7	7	7	7	1	3	7	−7	−21	−11	−17	5	−1	−35	−17	1	10
	9							4	28	14	14	4	4	7	2	−14	−22	−14	−11
	10													9	6	42	18	6	3
$\sum_{j=1}^{n} \{P_i(X_j)\}^2$		168	168	264	616	2184	264	60	2772	990	2002	468	1980	330	132	8580	2860	780	660
	λ	2	1	$\frac{2}{3}$	$\frac{7}{12}$	$\frac{7}{10}$	$\frac{11}{60}$	1	3	$\frac{5}{6}$	$\frac{7}{12}$	$\frac{3}{20}$	$\frac{11}{60}$	2	$\frac{1}{2}$	$\frac{5}{3}$	$\frac{5}{12}$	$\frac{1}{10}$	$\frac{11}{240}$

[a]Adaptada con permiso de *Biometrika Tables for Statisticians*, vol. 1, 3a. ed., E.S. Pearson y H.O. Hartley, Cambridge University Press, Cambridge

XI. Números aleatorios[a]

10480	15011	01536	02011	87647	91646	69179	14194	62590
22368	46573	25595	85393	30995	89198	27982	53402	93965
24130	48360	22527	97265	76393	64809	15179	24830	49340
42167	93093	06243	61680	07856	16376	39440	53537	71341
37570	39975	81837	16656	06121	91782	60468	81305	49684
77921	06907	11008	42751	27756	53498	18602	70659	90655
99562	72905	56420	69994	98872	31016	71194	18738	44013
96301	91977	05463	07972	18876	20922	94595	56869	69014
89579	14342	63661	10281	17453	18103	57740	84378	25331
85475	36857	53342	53988	53060	59533	38867	62300	08158
28918	69578	88231	33276	70997	79936	56865	05859	90106
63553	40961	48235	03427	49626	69445	18663	72695	52180
09429	93969	52636	92737	88974	33488	36320	17617	30015
10365	61129	87529	85689	48237	52267	67689	93394	01511
07119	97336	71048	08178	77233	13976	47564	81056	97735
51085	12765	51821	51259	77452	16308	60756	92144	49442
02368	21382	52404	60268	89368	19885	55322	44819	01188
01011	54092	33362	94904	31273	04146	18594	29852	71585
52162	53916	46369	58586	23216	14513	83149	98736	23495
07056	97628	33787	09998	42698	06691	76988	13602	51851
48663	91245	85828	14346	09172	30168	90229	04734	59193
54164	58492	22421	74103	47070	25306	76468	26384	58151
32639	32363	05597	24200	13363	38005	94342	28728	35806
29334	27001	87637	87308	58731	00256	45834	15398	46557
02488	33062	28834	07351	19731	92420	60952	61280	50001
81525	72295	04839	96423	24878	82651	66566	14778	76797
29676	20591	68086	26432	46901	20849	89768	81536	86645
00742	57392	39064	66432	84673	40027	32832	61362	98947
05366	04213	25669	26422	44407	44048	37937	63904	45766
91921	26418	64117	94305	26766	25940	39972	22209	71500
00582	04711	87917	77341	42206	35126	74087	99547	81817
00725	69884	62797	56170	86324	88072	76222	36086	84637
69011	65795	95876	55293	18988	27354	26575	08625	40801
25976	57948	29888	88604	67917	48708	18912	82271	65424
09763	83473	73577	12908	30883	18317	28290	35797	05998
91567	42595	27958	30134	04024	86385	29880	99730	55536
17955	56349	90999	49127	20044	59931	06115	20542	18059
46503	18584	18845	49618	02304	51038	20655	58727	28168
92157	89634	94824	78171	84610	82834	09922	25417	44137
14577	62765	35605	81263	39667	47358	56873	56307	67607

[a]Reproducida con permiso de *Probability and Statistics in Engineering and Management Science*, 3a. ed., W.W. Hines y D.C. Montgomery, Wiley, Nueva York

XII. Relaciones de alias para diseños factoriales fraccionados 2^{k-p} con $k \leq 15$ y $n \leq 64$

Diseños con 3 factores

a) 2^{3-1}; fracción 1/2 de 3 factores en 4 corridas — **Resolución III**

Generadores del diseño

$C = AB$

Relación de definición: $I = ABC$

Alias

$A = BC$

$B = AC$

$C = AB$

Diseños con 4 factores

b) 2^{4-1}; fracción 1/2 de 4 factores en 8 corridas — **Resolución IV**

Generadores del diseño

$D = ABC$

Relación de definición: $I = ABCD$

Alias

$A = BCD$

$B = ACD$

$C = ABD$

$D = ABC$

$AB = CD$

$AC = BD$

$AD = BC$

Diseños con 5 factores

c) 2^{5-2}; fracción 1/4 de 5 factores en 8 corridas — **Resolución III**

Generadores del diseño

$D = AB \quad E = AC$

Relación de definición: $I = ABD = ACE = BCDE$

Alias

$A = BD = CE$

$B = AD = CDE$

$C = AE = BDE$

$D = AB = BCE$

$E = AC = BCD$

$BC = DE = ACD = ABE$

$CD = BE = ABC = ADE$

d) 2^{5-1}; fracción 1/2 de 5 factores en 16 corridas — **Resolución V**

Generadores del diseño

$E = ABCD$

Relación de definición: $I = ABCDE$

Alias

Cada efecto principal es alias de una sola interacción de 4 factores

$AB = CDE \quad BD = ACE$

$AC = BDE \quad BE = ACD$

$AD = BCE \quad CD = ABE$

$AE = BCD \quad CE = ABD$

$BC = ADE \quad DE = ABC$

2 bloques de 8:

XII. Relaciones de alias para diseños factoriales fraccionados 2^{k-p} con $k \leq 15$ y $n \leq 64$ (*continuación*)

Diseños con 6 factores

e) 2^{6-3}; fracción 1/8 de 6 factores en 8 corridas — **Resolución III**

Generadores del diseño

$D = AB \quad E = AC \quad F = BC$

Relación de definición: $I = ABD = ACE = BCDE = BCF = ACDF = ABEF = DEF$

Alias

$A = BD = CE = CDF = BEF$

$B = AD = CF = CDE = AEF$

$C = AE = BF = BDE = ADF$

$D = AB = EF = BCE = ACF$

$E = AC = DF = BCD = ABF$

$F = BC = DE = ACD = ABE$

$CD = BE = AF = ABC = ADE = BDF = CEF$

f) 2^{6-2}; fracción 1/4 de 6 factores en 16 corridas — **Resolución IV**

Generadores del diseño

$E = ABC \quad F = BCD$

Relación de definición: $I = ABCE = BCDF = ADEF$

Alias

$A = BCE = DEF$

$B = ACE = CDF$

$C = ABE = BDF$

$D = BCF = AEF$

$E = ABC = ADF$

$F = BCD = ADE$

$ABD = CDE = ACF = BEF$

$ACD = BDE = ABF = CEF$

$AB = CE$

$AC = BE$

$AD = EF$

$AE = BC = DF$

$AF = DE$

$BD = CF$

$BF = CD$

2 bloques de 8: $ABD = CDE = ACF = BEF$

g) 2^{6-1}; fracción 1/2 de 6 factores en 32 corridas **Resolución VI**

Generadores del diseño

$F = ABCDE$

Relación de definición: $I = ABCDEF$

Alias

Cada efecto principal es alias de una sola interacción de 5 factores.
Cada interacción de 2 factores es alias de una sola interacción de 4 factores

$ABC = DEF \quad ACE = BDF$
$ABD = CEF \quad ACF = BDE$
$ABE = CDF \quad ADE = BCF$
$ABF = CDE \quad ADF = BCE$
$ACD = BEF \quad AEF = BCD$

2 bloques de 16: $ABC = DEF$ 4 bloques de 8: $AB = CDEF$, $ACD = BEF$, $AEF = BCD$

Diseños con 7 factores

h) 2^{7-4}; fracción 1/16 de 7 factores en 16 corridas **Resolución III**

Generadores del diseño

$D = AB \quad E = AC \quad F = BC \quad G = ABC$

Relación de definición: $I = ABD = ACE = BCDE = BCF = ACDF = ABEF = DEF = ABCG$
$= CDG = BEG = ADEG = AFG = BDFG = CEFG = ABCDEFG$

Alias

$A = BD = CE = FG \quad E = AC = DF = BG$
$B = AD = CF = EG \quad F = BC = DE = AG$
$C = AE = BF = DG \quad G = CD = BE = AF$
$D = AB = EF = CG$

XII. Relaciones de alias para diseños factoriales fraccionados 2^{k-p} con $k \leq 15$ y $n \leq 64$ (*continuación*)

i) 2^{7-3}; fracción 1/8 de 7 factores en 16 corridas — **Resolución IV**

Generadores del diseño

$E = ABC \quad F = BCD \quad G = ACD$

Relación de definición: $I = ABCE = BCDF = ADEF = ACDG = BDEG = ABFG = CEFG$

Alias

$A = BCE = DEF = CDG = BFG$	$AB = CE = FG$	$E = ABC = ADF = BDG = CFG$	$AF = DE = BG$
$B = ACE = CDF = DEG = AFG$	$AC = BE = DG$	$F = BCD = ADE = ABG = CEG$	$AG = CD = BF$
$C = ABE = BDF = ADG = EFG$	$AD = EF = CG$	$G = ACD = BDE = ABF = CEF$	$BD = CF = EG$
$D = BCF = AEF = ACG = BEG$	$AE = BC = DF$		

$ABD = CDE = ACF = BEF = BCG = AEG = DFG$

2 bloques de 8: $ABD = CDE = ACF = BEF = BCG = AEG = DFG$

j) 2^{7-2}; fracción 1/4 de 7 factores en 32 corridas — **Resolución IV**

Generadores del diseño

$F = ABCD \quad G = ABDE$

Relación de definición: $I = ABCDF = ABDEG = CEFG$

Alias

$A =$	$AB = CDF = DEG$	$BC = ADF$	$CE = FG$	$ACE = AFG$
$B =$	$AC = BDF$	$BD = ACF = AEG$	$CF = ABD = EG$	$ACG = AEF$
$C = EFG$	$AD = BCF = BEG$	$BE = ADG$	$CG = EF$	$BCE = BFG$
$D =$	$AE = BDG$	$BF = ACD$	$DE = ABG$	$BCG = BEF$
$E = CFG$	$AF = BCD$	$BG = ADE$	$DF = ABC$	$CDE = DFG$
$F = CEG$	$AG = BDE$	$CD = ABF$	$DG = ABE$	$CDG = DEF$
$G = CEF$				

2 bloques de 16: $ACE = AFG$ 4 bloques de 8: $ACE = AFG$
$BCE = BFG$
$AB = CDF = DEG$

k) 2^{7-1}; fracción 1/2 de 7 factores en 64 corridas — **Resolución VII**

<u>Generadores del diseño</u>

$G = ABCDEF$

Relación de definición: $I = ABCDEFG$

<u>Alias</u>

Cada efecto principal es alias de una sola interacción de 6 factores
Cada interacción de 2 factores es alias de una sola interacción de 5 factores
Cada interacción de 3 factores es alias de una sola interacción de 4 factores

2 bloques de 32: ABC 4 bloques de 16: ABC, CEF, CDG

Diseños con 8 factores

l) 2^{8-4}; fracción 1/16 de 8 factores en 16 corridas — **Resolución IV**

<u>Generadores del diseño</u>

$E = BCD \quad F = ACD \quad G = ABC \quad H = ABD$

Relación de definición: $I = BCDE = ACDF = ABEF = ABCG = ADEG = BDFG = CEFG = ABDH$
$= ACEH = BCFH = DEFH = CDGH = BEGH = AFGH = ABCDEFGH$

<u>Alias</u>

$A = CDF = BEF = BCG = DEG = BDH = CEH = FGH$ $AB = EF = CG = DH$
$B = CDE = AEF = ACG = DFG = ADH = CFH = EGH$ $AC = DF = BG = EH$
$C = BDE = ADF = ABG = EFG = AEH = BFH = DGH$ $AD = CF = EG = BH$
$D = BCE = ACF = AEG = BFG = ABH = EFH = CGH$ $AE = BF = DG = CH$
$E = BCD = ABF = ADG = CFG = ACH = DFH = BGH$ $AF = CD = BE = GH$
$F = ACD = ABE = BDG = CEG = BCH = DEH = AGH$ $AG = BC = DE = FH$
$G = ABC = ADE = BDF = CEF = CDH = BEH = AFH$ $AH = BD = CE = FG$
$H = ABD = ACE = BCF = DEF = CDG = BEG = AFG$

2 bloques de 8: $AB = EF = CG = DH$

XII. Relaciones de alias para diseños factoriales fraccionados 2^{k-p} con $k \leq 15$ y $n \leq 64$ (*continuación*)

m) 2^{8-3}; fracción 1/8 de 8 factores en 32 corridas **Resolución IV**

Generadores del diseño

$F = ABC \quad G = ABD \quad H = BCDE$

Relación de definición: $I = ABCF = ABDG = CDFG = BCDEH = ADEFH = ACEGH = BEFGH$

Alias

$A = BCF = BDG$	$AE = DFH = CGH$	$DE = BCH = AFH$
$B = ACF = ADG$	$AF = BC = DEH$	$DH = BCE = AEF$
$C = ABF = DFG$	$AG = BD = CEH$	$EF = ADH = BGH$
$D = ABG = CFG$	$AH = DEF = CEG$	$EG = ACH = BFH$
$E =$	$BE = CDH = FGH$	$EH = BCD = ADF = ACG = BFG$
$F = ABC = CDG$	$BH = CDE = EFG$	$FH = ADE = BEG$
$G = ABD = CDF$	$CD = FG = BEH$	$GH = ACE = BEF$
$H =$	$CE = BDH = AGH$	$ABE = CEF = DEG$
$AB = CF = DG$	$CG = DF = AEH$	$ABH = CFH = DGH$
$AC = BF = EGH$	$CH = BDE = AEG$	$ACD = BDF = BCG = AFG$
$AD = BG = EFH$		

2 bloques de 16: $ABE = CEF = DEG$

4 bloques de 8: $ABE = CEF = DEG$
$ABH = CFH = DGH$
$EH = BCD = ADF = ACG = BFG$

n) 2^{8-2}; fracción 1/4 de 8 factores en 64 corridas	**Resolución V**

<u>Generadores del diseño</u>

$G = ABCD \quad H = ABEF$

Relación de definición: $I = ABCDG = ABEFH = CDEFGH$

<u>Alias</u>

$AB = CDG = EFH$	$BG = ACD$	$EF = ABH$	$ADH =$	$BFG =$
$AC = BDG$	$BH = AEF$	$EG =$	$AEG =$	$BGH =$
$AD = BCG$	$CD = ABG$	$EH = ABF$	$AFG =$	$CDE = FGH$
$AE = BFH$	$CE =$	$FG =$	$AGH =$	$CDF = EGH$
$AF = BEH$	$CF =$	$FH = ABE$	$BCE =$	$CDH = EFG$
$AG = BCD$	$CG = ABD$	$GH =$	$BCF =$	$CEF = DGH$
$AH = BEF$	$CH =$	$ACE =$	$BCH =$	$CEG = DFH$
$BC = ADG$	$DE =$	$ACF =$	$BDE =$	$CEH = DFG$
$BD = ACG$	$DF =$	$ACH =$	$BDF =$	$CFG = DEH$
$BE = AFH$	$DG = ABC$	$ADE =$	$BDH =$	$CFH = DEG$
$BF = AEH$	$DH =$	$ADF =$	$BEG =$	$CGH = DEF$

2 bloques de 32: $CDE = FGH$ 4 bloques de 16: $CDE = FGH$
ACF
BDH

XII. Relaciones de alias para diseños factoriales fraccionados 2^{k-p} con $k \leq 15$ y $n \leq 64$ (*continuación*)

Diseños con 9 factores

o) 2^{9-5}; fracción 1/32 de 9 factores en 16 corridas — **Resolución III**

Generadores del diseño

$E = ABC \quad F = BCD \quad G = ACD \quad H = ABD \quad J = ABCD$

Relación de definición: $I = ABCE = BCDF = ADEF = ACDG = BDEG = ABFG = CEFG = ABDH$
$= CDEH = ACFH = BEFH = BCGH = AEGH = DFGH = ABCDEFGH = ABCDJ$
$= DEJ = AFJ = BCEFJ = BGJ = ACEGJ = CDFGJ = ABDEFGJ = CHJ$
$= ABEHJ = BDFHJ = ABCDEFHJ = ADGHJ = BCDEFGHJ = ABCFGHJ = EFGHJ$

Alias

$A = FJ$
$B = GJ$
$C = HJ$
$D = EJ$
$E = DJ$
$F = AJ$
$G = BJ$
$H = CJ$
$J = DE = AF = BG = CH$
$AB = CE = FG = DH$
$AC = BE = DG = FH$
$AD = EF = CG = BH$
$AE = BC = DF = GH$
$AG = CD = BF = EH$
$AH = BD = CF = EG$

2 bloques de 8: $AB = CE = FG = DH$

p) 2^{9-4}; fracción 1/16 de 9 factores
en 32 corridas

Resolución IV

<u>Generadores del diseño</u>

$F = BCDE \quad G = ACDE \quad H = ABDE \quad J = ABCE$

Relación de definición: $I = BCDEF = ACDEG = ABFG = ABDEH = ACFH = BCGH = DEFGH = ABCEJ$
$= ADFJ = BDGJ = CEFGJ = CDHJ = BEFHJ = AEGHJ = ABCDFGHJ$

<u>Alias</u>

$A = BFG = CFH = DFJ$	$AD = CEG = BEH = FJ$	$BJ = ACE = DG = EFH$
$B = AFG = CGH = DGJ$	$AE = CDG = BDH = BCJ = GHJ$	$CD = BEF = AEG = HJ$
$C = AFH = BGH = DHJ$	$AF = BG = CH = DJ$	$CE = BDF = ADG = ABJ = FGJ$
$D = AFJ = BGJ = CHJ$	$AG = CDE = BF = EHJ$	$CJ = ABE = EFG = DH$
$E =$	$AH = BDE = CF = EGJ$	$DE = BCF = ACG = ABH = FGH$
$F = ABG = ACH = ADJ$	$AJ = BCE = DF = EGH$	$EF = BCD = DGH = CGJ = BHJ$
$G = ABF = BCH = BDJ$	$BC = DEF = GH = AEJ$	$EG = ACD = DFH = CFJ = AHJ$
$H = ACF = BCG = CDJ$	$BD = CEF = AEH = GJ$	$EH = ABD = DFG = BFJ = AGJ$
$J = ADF = BDG = CDH$	$BE = CDF = ADH = ACJ = FHJ$	$EJ = ABC = CFG = BFH = AGH$
$AB = FG = DEH = CEJ$	$BH = ADE = CG = EFJ$	$AEF = BEG = CEH = DEJ$
$AC = DEG = FH = BEJ$		

2 bloques de 16: $AEF = BEG = CEH = DEJ$

4 bloques de 8: $AEF = BEG = CEH = DEJ$
$AB = FG = DEH = CEJ$
$CD = BEF = AEG = HJ$

XII. Relaciones de alias para diseños factoriales fraccionados 2^{k-p} con $k \leq 15$ y $n \leq 64$ (*continuación*)

Resolución IV

q) **2^{9-3}; fracción 1/8 de 9 factores en 64 corridas**

Generadores del diseño

$G = ABCD \quad H = ACEF \quad J = CDEF$

Relación de definición: $I = ABCDG = ACEFH = BDEFGH = CDEFJ = ABEFGJ = ADHJ = BCGHJ$

Alias

$A = DHJ$	$AC = BDG = EFH$	$BF =$
$B =$	$AD = BCG = HJ$	$BG = ACD = CHJ$
$C =$	$AE = CFH$	$BH = CGJ$
$D = AHJ$	$AF = CEH$	$BJ = CGH$
$E =$	$AG = BCD$	$CD = ABG = EFJ$
$F =$	$AH = CEF = DJ$	$CE = AFH = DFJ$
$G =$	$AJ = DH$	$CF = AEH = DEJ$
$H = ADJ$	$BC = ADG = GHJ$	$CG = ABD = BHJ$
$J = ADH$	$BD = ACG$	$CH = AEF = BGJ$
$AB = CDG$	$BE =$	$CJ = DEF = BGH$
$DE = CFJ$	$GJ = BCH$	$AFJ = BEG = DFH$
$DF = CEJ$	$ABE = FGJ$	$AGH = DGJ$
$DG = ABC$	$ABF = EGJ$	$AGJ = BEF = DGH$
$EF = ACH = CDJ$	$ABH = BDJ$	$BCE =$
$EG =$	$ABJ = EFG = BDH$	$BCF =$
$EH = ACF$	$ACJ = CDH$	$BDE = FGH$
$EJ = CDF$	$ADE = EHJ$	$BDF = EGH$
$FG =$	$ADF = FHJ$	$BEH = DFG$
$FH = ACE$	$AEG = BFJ$	$BFG = DEG$
$FJ = CDE$	$AEJ = BFG = DEH$	$CEG =$
$GH = BCJ$	$AFG = BEJ$	$CFG =$

2 bloques de 32: CFG 4 bloques de 16: CFG

$AGJ = BEF = DGH$

$ADE = EHJ$

Diseños con 10 factores

r) 2^{10-6}; fracción 1/64 de 10 factores en 16 corridas

Resolución III

Generadores del diseño

$E = ABC \quad F = BCD \quad G = ACD \quad H = ABD \quad J = ABCD \quad K = AB$

Relación de definición: $I = ABCE = BCDF = ADEF = ACDG = BDEG = ABFG = CEFG = ABDH$
$= CDEH = ACFH = BEFH = BCGH = AEGH = DFGH = ABCDEFGH = ABCDJ$
$= DEJ = AFJ = BCEFJ = BGJ = ACEGJ = CDFGI = ABDEFGJ = CHJ$
$= ABEHJ = BDFHJ = ACDEFHJ = ADGHJ = BCDEGHJ = ABCFGHJ = EFGHJ = ABK$
$= CEK = ACDFK = BDEFK = BCDGK = ADEGK = FGK = ABCEFGK = DHK$
$= ABCDEHK = BCFHK = AEFHK = ACGKH = BEGHK = ABDFGHK = CDEFGHK = CDJK$
$= ABDEJK = BFJK = ACEFJK = AGJK = BCEGJK = ABCDFGJK = DEFGJK = ABCHJK$
$= EHJK = ADFHJK = BCDEFHJK = BDGHJK = ACDEGHJK = CFGHJK = ABEFGHJK$

Alias

$A = FJ = BK$	$J = DE = AF = BG = CH$
$B = GJ = AK$	$K = AB = CE = FG = DH$
$C = HJ = EK$	$AC = BE = DG = FH$
$D = EJ = HK$	$AD = EF = CG = BH$
$E = DJ = CK$	$AE = BC = DF = GH$
$F = AJ = GK$	$AG = CD = BF = EH = JK$
$G = BJ = FK$	$AH = BD = CF = EG$
$H = CJ = DK$	

2 bloques de 8: $AG = CD = BF = EH = JK$

XII. Relaciones de alias para diseños factoriales fraccionados 2^{k-p} con $k \leq 15$ y $n \leq 64$ (*continuación*)

s) 2^{10-5}; fracción 1/32 de 10 factores en 32 corridas **Resolución IV**

Generadores del diseño

$F = ABCD \quad G = ABCE \quad H = ABDE \quad J = ACDE \quad K = BCDE$

Relación de definición: $I = ABCDF = ABCEG = DEFG = ABDEH = CEFH = CDGH = ABFGH = ACDEJ$
$= BEFJ = BDGJ = ACFGJ = BCHJ = ADFHJ = AEGHJ = BCDEFGHJ = BCDEK$
$= AEFK = ADGK = BCFGK = ACHK = BDFHK = BEGHK = ACDEFGHK = ABJK$
$= CDFJK = CEGJK = ABDEFGJK = DEHJK = ABCEFHJK = ABCDGHJK = FGHJK$

Alias

$A = EFK = DGK = CHK = BJK$
$B = EFJ = DGJ = CHJ = AJK$
$C = EFH = DGH = BHJ = AHK$
$D = EFG = CGH = BGJ = AGK$
$E = DFG = CFH = BFJ = AFK$
$F = DEG = CEH = BEJ = AEK$
$G = DEF = CDH = BDJ = ADK$
$H = CEF = CDG = BCJ = ACK$
$J = BEF = BDG = BCH = ABK$
$K = AEF = ADG = ACH = ABJ$
$AB = CDF = CEG = DEH = FGH = JK$
$AC = BDF = BEG = DEJ = FGJ = HK$
$AD = BCF = BEH = CEJ = FHJ = GK$
$AE = BCG = BDH = CDJ = GHJ = FK$
$AF = BCD = BGH = CGJ = DHJ = EK$
$AG = BCE = BFH = CFJ = EHJ = DK$
$AH = BDE = BFG = DFJ = EGJ = CK$
$AJ = CDE = CFG = DFH = EGH = BK$
$AK = EF = DG = CH = BJ$
$BC = ADF = AEG = HJ = DEK = FGK$
$BD = ACF = AEH = GJ = CEK = FHK$
$BE = ACG = ADH = FJ = CDK = GHK$
$BF = ACD = AGH = EJ = CGK = DHK$
$BG = ACE = AFH = DJ = CFK = EHK$
$BH = ADE = AFG = CJ = DFK = EGK$
$CD = ABF = GH = AEJ = BEK = FJK$
$CE = ABG = FH = ADJ = BDK = GJK$
$CF = ABD = EH = AGJ = BGK = DJK$
$CG = ABE = DH = AFJ = BFK = EJK$
$DE = FG = ABH = ACJ = BCK = HJK$
$DF = ABC = EG = AHJ = BHK = CJK$

2 bloques de 16: $AK = EF = DG = CH = BJ$
4 bloques de 8: $AK = EF = DG = CH = BJ$
$AJ = CDE = CFG = DFH = EGH = BK$
$AB = CDF = CEG = DEH = FGH = JK$

t) 2^{10-4}; fracción 1/16 de 10 factores en 64 corridas **Resolución IV**

Generadores del diseño

$G = BCDF \quad H = ACDF \quad J = ABDE \quad K = ABCE$

Relación de definición: $I = BCDFG = ACDFH = ABGH = ABDEJ = ACEFGJ = BCEFHJ = DEGHJ = ABCEK$
$= ADEFGK = BDEFHK = CEGHK = CDJK = BFGJK = AFHJK = ABCDGHJK$

Alias

$A = BGH$	$AD = CFH = BEJ$	$BK = ACE = FGJ$
$B = AGH$	$AE = BDJ = BCK$	$CD = BFG = AFH = JK$
$C = DJK$	$AF = CDH = HJK$	$CE = ABK = GHK$
$D = CJK$	$AG = BH$	$CF = BDG = ADH$
$E =$	$AH = CDF = BG = FJK$	$CG = BDF = EHK$
$F =$	$AJ = BDE = FHK$	$CH = ADF = EGK$
$G = ABH$	$AK = BCE = FHJ$	$CJ = DK$
$H = ABG$	$BC = DFG = AEK$	$CK = ABE = EGH = DJ$
$J = CDK$	$BD = CFG = AEJ$	$DE = ABJ = GHJ$
$K = CDJ$	$BE = ADJ = ACK$	$DF = BCG = ACH$
$AB = GH = DEJ = CEK$	$BF = CDG = GJK$	$DG = BCF = EHJ$
$AC = DFH = BEK$	$BJ = ADE = FGK$	$DH = ACF = EGJ$
$EF =$	$GJ = DEH = BFK$	$AEG = BEH = CFJ = DFK$
$EG = DHJ = CHK$	$GK = CEH = BFJ$	$AEH = BEG$
$EH = DGJ = CGK$	$HJ = DEG = AFK$	$AFG = BFH = CEJ = DEK$
$EJ = ABD = DGH$	$HK = CEG = AFJ$	$AGJ = CEF = BHJ$
$EK = ABC = CGH$	$ABF = FGH$	$AGK = DEF = BHK$
$FG = BCD = BJK$	$ACG = BCH = EFJ$	$BCJ = EFH = BDK$
$FH = ACD = AJK$	$ACJ = EFG = ADK$	$BEF = CHJ = DHK$
$FJ = BGK = AHK$	$ADG = BDH = EFK$	$CDE = EJK$
$FK = BGJ = AHJ$	$AEF = CGJ = DGK$	$CFK = DFJ$

2 bloques de 32: $AGJ = CEF = BHJ$ 4 bloques de 16: $AGJ = CEF = BHJ$
$AGK = DEF = BHK$
$CD = BFG = AFH = JK$

XII. Relaciones de alias para diseños factoriales fraccionados 2^{k-p} con $k \leq 15$ y $n \leq 64$ (*continuación*)

Diseños con 11 factores

u) 2^{11-7}; fracción 1/128 de 11 factores en 16 corridas — **Resolución III**

Generadores del diseño

$E = ABC \quad F = BCD \quad G = ACD \quad H = ABD \quad J = ABCD \quad K = AB \quad L = AC$

Relación de definición: $I = ABCE = BCDF = ADEF = ACDG = BDEG = ABFG = CEFG = ABDH$
$= CDEH = ACFH = BEFH = BCGH = AEGH = DFGH = ABCDEFGH = ABCDJ$
$= DEJ = AFJ = BCEFJ = BGJ = ACEGJ = CDFGJ = ABDEFGJ = CHJ$
$= ABEHJ = BDFHJ = ACDEFHJ = ADGHJ = BCDEGHJ = ABCFGHJ = EFGHJ = ABK$
$= CEK = ACDFK = BDEFK = BCDGK = ADEGK = FGK = ABCEFGK = DHK$
$= ABCDEHK = BCFHK = AEFHK = ACGHK = BEGHK = ABDFGHK = CDEFGHK = CDJK$
$= ABDEJK = BFJK = ACEFJK = AGJK = BCEGJK = ABCDFGJK = DEFGJK = ABCHJK$
$= EHJK = ADFHJK = BCDEFHJK = BDGHJK = ACDEGHJK = CFGHJK = ABEFGHJK = ACL$
$= BEL = ABDFL = CDEFL = DGL = ABCDEGL = BCFGL = AEFGL = BCDHL$
$= ADEHL = FHL = ABCEFHL = ABGHL = CEGHL = ACDFGHL = BDEFGHL = BDJL$
$= ACDEJL = CFJL = ABEFJL = ABCGJL = EGJL = ADFGJL = BCDEFGJL = AHJL$
$= BCEHJL = ABCDFHJL = DEFHJL = CDGHJL = ABDEGHJL = BFGHJL = ACEFGHJL = BCKL$
$= AEKL = DFKL = ABCDEFKL = ABDGKL = CDEGKL = ACFGKL = BEFGKL = ACDHKL$
$= BDEHKL = ABFHKL = CEFHKL = GHKL = ABCEGHKL = BCDFGHKL = ADEFGHKL = ADJKL$
$= BCDEJKL = ABCFJKL = EFJKL = CGJKL = ABEGJKL = BDFGJKL = ACDEFGJKL = BHJKL$
$= ACEHJKL = CDFHJKL = ABDEFHJKL = ABCDGHJKL = DEGHJKL = AFGHJKL = BCEFGHJKL$

Alias

$A = FJ = BK = CL$	$J = DE = AF = BG = CH$
$B = GJ = AK = EL$	$K = AB = CE = FG = DH$
$C = HJ = EK = AL$	$L = AC = BE = DG = FH$
$D = EJ = HK = GL$	$AD = EF = CG = BH$
$E = DJ = CK = BL$	$AE = BC = DF = GH = KL$
$F = AJ = GK = HL$	$AG = CD = BF = EH = JK$
$G = BJ = FK = DL$	$AH = BD = CF = EG = JL$
$H = CJ = DK = FL$	

2 bloques de 8: $AE = BC = DF = GH = KL$

v) 2^{11-6}; fracción 1/64 de 11 factores en 32 corridas **Resolución IV**

Generadores del diseño

$$F = ABC \quad G = BCD \quad H = CDE \quad J = ACD \quad K = ADE \quad L = BDE$$

Relación de definición: $I = ABCF = BCDG = ADFG = CDEH = ABDEFH = BEGH = ACEFGH = ACDJ = BDFJ = ABGJ = CFGJ$
$= AEHJ = BCEFHJ = ABCDEGHJ = DEFGHJ = ADEK = BCDEFK = ABCEGK = EFGK = ACHK = BFHK$
$= ABDGHK = CDFGHK = CEJK = ABEFJK = BDEGJK = ACDEFGJK = DHJK = ABCDFHJK = BCGHJK$
$= AFGHJK = BDEL = ACDEFL = CEGL = ABEFGL = BCHL = AFHL = DGHL = ABCDFGHL$
$= ABCEJL = EFJL = ADEGJL = BCDEFGJL = ABDHJL = CDFHJL = ACGHJL = BFGHJL = ABKL$
$= CFKL = ACDGKL = BDFGKL = ABCDEHKL = DEFHKL = AEGHKL = BCEFGHKL = BCDJKL$
$= ADFJKL = GJKL = ABCFGJKL = BEHJKL = ACEFHJKL = CDEGHJKL = ABDEFGHJKL$

Alias

$A = BCF = DFG = CDJ = BGJ = EHJ = DEK = CHK = FHL = BKL$
$B = ACF = CDG = EGH = DFJ = AGJ = FHK = DEL = CHL = AKL$
$C = ABF = BDG = DEH = ADJ = FGJ = AHK = EJK = EGL = BHL = FKL$
$D = BCG = AFG = CEH = ACJ = BFJ = AEK = HJK = BEL = GHL$
$E = CDH = BGH = AHJ = ADK = FGK = CJK = BDL = CGL = FJL$
$F = ABC = ADG = BDJ = CGJ = EGK = BHK = AHL = EJL = CKL$
$G = BCD = ADF = BEH = ABJ = CFJ = EFK = CEL = DHL = JKL$
$H = CDE = BEG = AEJ = ACK = BFK = DJK = BCL = AFL = DGL$
$J = ACD = BDF = ABG = CFG = AEH = CEK = DHK = EFL = GKL$
$K = ADE = EFG = ACH = BFH = CEJ = DHJ = ABL = CFL = GJL$
$L = BDE = CEG = BCH = AFH = DGH = EFJ = ABK = CFK = GJK$

$AB = CF = GJ = KL$	$AE = HJ = DK$	$AH = EJ = CK = FL$	$AL = FH = BK$	$BH = EG = CL = FK$
$AC = BF = DJ = HK$	$AF = BC = DG = HL$	$AJ = CD = BG = EH$	$BD = CG = FJ = EL$	$CE = DH = JK = GL$
$AD = FG = CJ = EK$	$AG = DF = BJ$	$AK = DE = CH = BL$	$BE = GH = DL$	$EF = GK = JL$

$ABD = CDF = ACG = BFG = EFH = BCJ = AFJ = DGJ = BEK = GHK = AEL = HJL = DKL$
$ABE = CEF = DFH = AGH = EGJ = BHJ = BDK = CGK = FJK = ADL = FGL = CJL = EKL$
$ABH = DEF = AEG = CFH = BEJ = GHJ = BCK = AFK = DGK = ACL = BFL = DJL = HKL$
$ACE = BEF = ADH = FGH = DEJ = CHJ = CDK = BGK = EHK = AJK = DFL = AGL = BJL$
$AEF = BCE = DEG = BDH = CGH = FHJ = DFK = AGK = BJK = CDL = BGL = EHL = AJL$

2 bloques de16: $AB = CF = GJ = KL$ 4 bloques de 8: $AB = CF = GJ = KL$
$AD = FG = CJ = EK$
$BD = CG = FJ = EL$

XII. Relaciones de alias para diseños factoriales fraccionados 2^{k-p} con $k \leq 15$ y $n \leq 64$ (*continuación*)

Diseños con 12 factores

w) 2^{12-8}; fracción 1/256 de 12 factores en 16 corridas **Resolución III**

Generadores del diseño

$E = ABC \quad F = ABD \quad G = ACD \quad H = BCD$
$J = ABCD \quad K = AB \quad L = AC \quad M = AD$

Alias

$A = HJ = BK = CL = DM$
$B = GJ = AK = EL = FM$
$C = FJ = EK = AL = GM$
$D = EJ = FK = GL = AM$
$E = DJ = CK = BL = HM$
$F = CJ = DK = HL = BM$
$G = BJ = HK = DL = CM$
$H = AJ = GK = FL = EM$
$J = DE = CF = BG = AH$
$K = AB = CE = DF = GH$
$L = AC = BE = DG = FH$
$M = AD = BF = CG = EH$
$AE = BC = FG = DH = KL = JM$
$AF = BD = EG = CH = JL = KM$
$AG = EF = CD = BH = JK = LM$

2 bloques de 8: $AE = BC = FG = DH = KL = JM$

Diseños con 13 factores

x) 2^{13-9}; fracción 1/512 de 13 factores en 16 corridas **Resolución III**

Generadores del diseño

$E = ABC \quad F = ABD \quad G = ACD \quad H = BCD$
$J = ABCD \quad K = AB \quad L = AC \quad M = AD \quad N = BC$

Alias

$A = HJ = BK = CL = DM = EN$
$B = GJ = AK = EL = FM = CN$
$C = FJ = EK = AL = GM = BN$
$D = EJ = FK = GL = AM = HN$
$E = DJ = CK = BL = HM = AN$
$F = CJ = DK = HL = BM = GN$
$G = BJ = HK = DL = CM = FN$
$H = AJ = GK = FL = EM = DN$
$J = DE = CF = BG = AH = MN$
$K = AB = CE = DF = GH = LN$
$L = AC = BE = DG = FH = KN$
$M = AD = BF = CG = EH = JN$
$N = BC = AE = FG = DH = KL = JM$
$AF = BD = EG = CH = JL = KM$
$AG = EF = CD = BH = JK = LM$

2 bloques de 8: $AF = BD = EG = CH = JL = KM$

XII. Relaciones de alias para diseños factoriales fraccionados 2^{k-p} con $k \leq 15$ y $n \leq 64$ (*continuación*)

Diseños con 14 factores

y) 2^{14-10}; fracción 1/1024 de 14 factores en 16 corridas — **Resolución III**

Generadores del diseño

$E = ABC \quad F = ABD \quad G = ACD \quad H = BCD \quad J = ABCD$
$K = AB \quad L = AC \quad M = AD \quad N = BC \quad O = BD$

Alias

$A = HJ = BK = CL = DM = EN = FO$
$B = GJ = AK = EL = FM = CN = DO$
$C = FJ = EK = AL = GM = BN = HO$
$D = EJ = FK = GL = AM = HN = BO$
$E = DJ = CK = BL = HM = AN = GO$
$F = CJ = DK = HL = BM = GN = AO$
$G = BJ = HK = DL = CM = FN = EO$
$H = AJ = GK = FL = EM = DN = CO$
$J = DE = CF = BG = AH = MN = LO$
$K = AB = CE = DF = GH = LN = MO$
$L = AC = BE = DG = FH = KN = JO$
$M = AD = BF = CG = EH = JN = KO$
$N = BC = AE = FG = DH = KL = JM$
$O = BD = AF = EG = CH = JL = KM$
$AG = EF = CD = BH = JK = LM = NO$

2 bloques de 8: $AG = EF = CD = BH = JK = LM = NO$

Diseños con 15 factores

z) 2^{15-11}; fracción 1/2048 de 15 factores en 16 corridas — **Resolución III**

Generadores del diseño

$E = ABC \quad F = ABD \quad G = ACD \quad H = BCD \quad J = ABCD$
$K = AB \quad L = AC \quad M = AD \quad N = BC \quad O = BD \quad P = CD$

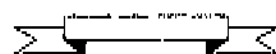
Alias

$A = HJ = BK = CL = DM = EN = FO = GP$
$B = GJ = AK = EL = FM = CN = DO = HP$
$C = FJ = EK = AL = GM = BN = HO = DP$
$D = EJ = FK = GL = AM = HN = BO = CP$
$E = DJ = CK = BL = HM = AN = GO = FP$
$F = CJ = DK = HL = BM = GN = AO = EP$
$G = BJ = HK = DL = CM = FN = EO = AP$
$H = AJ = GK = FL = EM = DN = CO = BP$
$J = DE = CF = BG = AH = MN = LO = KP$
$K = AB = CE = DF = GH = LN = MO = JP$
$L = AC = BE = DG = FH = KN = JO = MP$
$M = AD = BF = CG = EH = JN = KO = LP$
$N = BC = AE = FG = DH = KL = JM = OP$
$O = BD = AF = EG = CH = JL = KM = NP$
$P = CD = EF = AG = BH = JK = LM = NO$

Tabla XIII. Glosario para el uso de *Design Expert*

Actual value	Valor real
Adeq precision	Precisión adecuada
Adj MS (adjusted mean square)	Cuadrado medio ajustado
Adj R-squared	*R* cuadrada ajustada
Adj SS (adjusted sum of squares)	Suma de cuadrados ajustada (SS ajustada)
Analysis of variance	Análisis de varianza
Coefficient estimate	Estimación del coeficiente
C.V. (coefficient of variation)	Coeficiente de variación
CI (confidence interval)	Intervalo de confianza
Cook's distance	Distancia de Cook
Cor total (corrected total)	Total corregido. Suma de los valores de respuesta corregidos por la media. Se llama comúnmente suma de cuadrados totales
DF (degrees of freedom)	Grados de libertad
DDF (denominator degrees of freedom)	Grados de libertad del denominador
Diagnostic case statistics	Estadísticos de diagnóstico del caso
Error term	Término del error
Estimated mean	Media estimada
Expected mean square	Cuadrado medio esperado
Lack of fit	Falta de ajuste
Least squares means for strength	Medias de mínimos cuadrados de la resistencia
Leverage	Acción de palanca. Potencial de un punto del diseño para influir en los coeficientes del ajuste del modelo. Deben evitarse valores cercanos a uno
Mean	Media
Mean difference	Diferencia media
Mean square	Cuadrado medio
Means for covariates	Medias de las covariables
NDF (numerator degrees of freedom)	Grados de libertad del numerador
Operator fixed	Operador fijo
Operator random	Operador aleatorio
Outlier t	Punto atípico *t*
Pred R-squared	*R* cuadrada predicha
Predicted value	Valor predicho
PRESS (Prediction error sum of squares)	Suma de cuadrados del error de predicción
Pure error	Error puro
R-squared	R cuadrada
Root MSE (root mean square error)	Raíz cuadrada del cuadrado medio del error
Seq SS (sequential sum of squares)	Suma de cuadrados secuencial
Std. Dev. (standard deviation)	Desviación estándar
Standard error (SE)	Error estándar
SE of difference (standard error of difference)	Error estándard de la diferencia
Std. error of estimation	Error estándar de la estimación
SE mean (standard error of the mean)	Error estándar de la media
Standard order	Orden estándar
Student residual	Residual de Student
Sum of squares	Suma de cuadrados
Variance component	Componente de la varianza

Índice



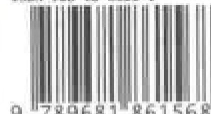
La edición, composición, diseño e impresión de esta obra fueron realizados
bajo la supervisión de GRUPO NORIEGA EDITORES.
Balderas 95, Col. Centro. México, D.F. C.P. 06040
2225750000704658DP9233IE

Cómo realizar una experimentación industrial óptima. Esta obra para ingenieros, científicos y expertos en probabilidad proporciona el enfoque más eficaz para aprender a diseñar, conducir y analizar experimentos que optimicen la calidad y el desempeño de productos y proceso. En esta nueva edición completamente revisada y actualizada, el autor trata con mayor detenimiento los diseños factoriales y factoriales fraccionados, y presenta nuevas técnicas de análisis, como el modelo lineal generalizado. Asimismo, ofrece un estudio más amplio de los experimentos con factores aleatorios, la metodología de superficies de respuesta, los experimentos con mezclas y los métodos para estudios de robustez de procesos.

El libro también ilustra dos de las herramientas de software más poderosas hoy en día para el diseño experimental: Design Expert® y Minitab®. A lo largo del texto se incluyen salidas de estos programas, junto con una explicación detallada de la forma en que se están usando las computadoras para el análisis y el diseño de experimentos.

* El sitio del libro en la *web* ofrece recursos adicionales para estudiantes y profesores.

ÁREA: ING. INDUSTRIAL
ISBN 968-18-6156-6



e-mail: limusa@noriega.com.mx
www.noriega.com.mx